

Psychological Review

EDITED BY

CARROLL C. PRATT
PRINCETON UNIVERSITY

CONTENTS

Ontogenesis of Emotional Behavior.....	W. A. BOUSFIELD, & W. D. ORRISON	1
Primary Stimulus Generalization: A Neurophysiological View.....	JOSEPH WOLPE	8
A Preface to a Psychological Analysis of the Self.....	THEODORE R. SARBIN	11
Structural Rigidity in Relation to Learning Theory and Clinical Psychol- ogy.....	RAYMOND B. CATTELL, & ALVIN E. WINDER	23
The Process Need.....	IRVING MALTZMAN	40
On the Nature of Inhibition in the Cerebral Cortex.....	MAX HAMILTON	49
Responses in Absence of the Acquisi- tion Motive.....	WILSE B. WEBB	54
Behavior and the Psychophysical Meth- ods: An Analysis of Some Recent Experiments.....	C. H. GRAHAM	62
The Interrupted Task Method in Stud- ies of Selective Recall: A Reevaluation of Some Recent Experiments.....	THELMA G. ALPER	71
The Nature of the Response in Dis- crimination Learning.....	KENNETH W. SPENCE	89

PUBLISHED BI-MONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

The *Psychological Review* is devoted primarily to articles in the field of general and theoretical psychology. This area is obviously difficult to define and delimit, but in view of the large number of manuscripts sent to the editor on all kinds of topics an attempt has to be made to draw the line somewhere.

Ordinarily manuscripts that run to more than about 7500 words are not accepted. This policy is followed partly in an effort to reduce lag of publication and partly from the conviction that brevity which is not inconsistent with clarity is the best way to present an argument.

If an author is prepared to pay for the cost of printing his article, he may arrange for earlier publication without thereby postponing the appearance of manuscripts by other contributors.

Tables, footnotes and references as well as text of manuscripts should be typed double-spaced throughout.

PUBLISHED BI-MONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.
\$5.50 volume \$1.00 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (4-3), Section 3440, P. L. & R. of 1948, authorized Jan. 8, 1948

THE PSYCHOLOGICAL REVIEW

ONTOGENESIS OF EMOTIONAL BEHAVIOR

BY W. A. BOUSFIELD AND W. D. ORBISON

University of Connecticut

Descriptions of children's emotional behavior appearing in the psychological literature contain statements to the effect that early emotional response patterns tend to be relatively brief, intense, frequent, and undifferentiated in character. Furthermore, the range of effective stimuli for early emotional responses is comparatively narrow. Jersild has effectively summarized observational studies indicating these characteristics and their changes with advancing age. In his words, "There is a trend away from gross and explosive overt responses toward more subtle and indirect expressions, and from momentary outbursts to more prolonged emotional states" (14, p. 759). Various psychological assumptions have been made to account for the reported findings. The child's span of attention is short and easily diverted. He lacks emotional control. His uninhibited expression quickly discharges emotional tension. In Lewinian terms we may say that the barriers between the regions of the child's personality are relatively permeable. Hence emotional tension spreads easily and is soon dissipated. Such explanations, we would submit, may have limited value. Certainly such concepts as emotional tension and emotional control are difficult to define. The purpose of this paper is to indicate that the major findings are

corollaries of the following two physiological assumptions:

1. *The infant is essentially a precorticate organism, and the establishment of cortical control, especially in the frontal lobes, is not achieved until adulthood.*

2. *The infant is relatively lacking in endocrine products which sustain some of the physiological responses to stress.*

With regard to the first assumption we shall consider first, the evidence showing that the cortex functions in emotional behavior; second, the evidence showing that the cortex, and especially the frontal lobes, has a long developmental history. For the purpose of systematic clarity we shall follow Hebb (12) in distinguishing between the terms *emotion* and *emotional behavior*. Hebb (12, p. 148) defines emotion as a construct denoting "the neural process that is inferred from and causes emotional behavior." Our assumption of the precorticate status of the infant necessarily implies this construct though we are not committed to the definition of specific neural mechanisms. In discussing emotional behavior we shall also follow Hebb in using the term to denote both emotional upset or disturbance and organized emotional responses such as those seen in aggression or flight.

If the newborn infant may be described as precorticate, it would follow

that it should resemble in several ways the thalamic organism, and that we should observe some similarities between early expressions of emotion in the human infant and the emotional behavior of decorticate animals. Experimental studies appear to provide a measure of confirmation to this expectation. The sham rage reactions in decorticate cats reported by Bard (2) were described as being relatively gross, automatic and intense. As would be expected, these reactions followed the application of directly impinging stimuli. That Harlow and Stagner (10) have questioned the propriety of the term *sham rage* to denote the emotional behavior of decorticate animals, and would substitute *excitement*, strengthens the suggested parallel, since Bridges (4) also prefers the term *excitement* to describe the undifferentiated precursor of specific emotional patterns in human infants.

The recent study by Bard and Mountcastle (3) has extended our knowledge of the functioning of the cortex in emotional expression. They have shown that cats can be made extremely placid by removing all of the cortex except two centers, namely, the amygdaloid complex and the transitional cortex of the midline. Removal of either of these two centers changes the cat into an extremely ferocious animal. Ferocity could also be induced by removing only the amygdaloid complex and the pyriform lobe, i.e., leaving the rest of the cortex intact. These latter animals differ from the thalamic animals in that their rage reactions are better directed and better timed.

Evidence that the cortex, and especially the frontal lobes, functions in emotional expression comes also from the studies on frontal lobotomies. Freeman and Watts (8), in describing the behavior of the typical patient after frontal lobotomy, report a variety of characteristics which would appear also

to apply to the early stages of emotional development. The responses of the operated patients to external impressions are immediate and sometimes vivid. Emotional behavior tends to be lacking in depth. The patients are possibly over-responsive. Their expressions of emotion appear to become purer, more uninhibited, and lacking in after-effects. These responses are of short duration, and transitions from one emotional state to another occur rapidly. The behavior of the operated patients shows a quality which is described as *childish*. Russell (18), in his analysis of the functions of the frontal lobes, states that simple emotional responses to afferent stimuli are most clearly observable in children, and these responses resemble those of the adult following severe frontal injury. He believes it is probable that one of the functions of the prefrontal lobes is involved in the development and control of emotional reactions. To this we may add Hebb's (12) relevant observation that as one goes down the phylogenetic scale it appears that there is a progressive shortening of the periods of emotional disturbance which follow momentary stimulation.

On the basis of the above-mentioned evidence it would seem reasonable to suppose that the cortex, and especially the frontal lobes, is somehow involved in the inhibition, instigation and sustaining of emotional reactions.

The ontogenetic pattern of emotional behavior is here assumed to result from the fact that the infant is essentially precorticate at birth and that the establishment of maximum cortical control (or influence) requires a long time. Textbooks portraying neurological development, e.g., Ford (7), contain statements to the effect that while the cranial and spinal segments are capable of relatively adequate functional activity at birth, the cerebral cortex appears to

exert little influence on the segmental apparatus. Histological examination of cortical cells reveals a relative lack of medulation and a deficiency in Nissl granules. According to Flechsig (6), the prefrontal lobes, especially the areas involved in prefrontal lobotomy, are among the last of the cortical areas to become myelinated.

Gibbs and Gibbs (9), in summarizing the data on the developmental changes in the EEG, report that the newborn infant shows no steady frequencies. The patterns include slow waves of $\frac{1}{2}$ to 2 cycles per second upon which are superimposed fast waves of 20 to 50 per second. The parietal areas show the greatest activity. With increasing age there is a tendency for the slow waves to become faster, but the rate of increase varies in different cortical areas. By the age of six months the activity of the occipital areas exceeds that of the parietal as well as that of the frontal regions. At the ninth year adult frequencies are found in the occipital areas, but the parietal and frontal regions still show slower frequencies than those of the adult. Scattered forms of 7-per-second waves are found in the parietal areas during the ages of 12 to 15. From the ages of 14 to 17 waves with frequencies of 5 to 7 per second are common in the frontal areas. It is not until the age of 19 that adult patterns are found in all cortical areas. From the point of view of this paper it is significant that the frontal areas do not show adult characteristics until relatively late in ontogenetic development.

Assuming the validity of the foregoing analysis, we may say that the neural mechanisms permitting interaction between the hypothalamus and cortical areas, especially the frontal lobes, are incompletely established at birth. Furthermore, their maximum growth is reached relatively late in development. That the so-called higher

centers are capable of modifying emotional behavior appears evident from studies designed to test the consequences of the removal of cortical areas. We therefore conclude that the ontogenesis of emotional states is to a large extent dependent on the development of the frontal lobes. We should point out, however, that we cannot exclude areas other than the frontal lobes as agents in the control of emotional behavior. Hebb's (11) critical analysis leads him to conclude that cortical centers other than those in the frontal lobes are involved in emotional expression. Furthermore, centers other than the hypothalamus are involved in sham rage.

It is perhaps futile for us at this time to speculate on the specific neural mechanisms that constitute emotion. We may note, however, that Hebb (12) has proposed a central neurological theory of emotional disturbance which appears to account for most of the reported findings relative to emotion in children. The theory does have a weakness which Hebb is careful to point out, namely, that it does not account for the increase in the vigor of emotional states with increasing age. It is attractive to suppose that our second assumption of the child's deficiency in endocrine products provides a useful supplement. We shall now consider this assumption, with the realization that it is highly speculative. It is suggested, but far from proven, by the evidence available to us. We venture further to say that it has the merit of being testable, and research proving its validity or lack of validity would add significantly to our knowledge of the nature of emotion.

Our consideration will center on the adrenal glands, though it is likely that other glands such as the pituitary have a significant influence on the nature of emotion and the course of emotional development. Scammon (20) has plotted growth curves for various endocrine or-

gans on the basis of per cent of their total postnatal increments up to the age of 20. The development of the adrenals is peculiar. After birth they decrease rapidly in size until about the age of two when they are approximately 120 per cent below their birth weight. At this point gain in weight sets in rather abruptly. The gain is less rapid between the ages of five and eleven. From the ages of eleven to twenty the curve rises smoothly and steeply with a positive acceleration. The birth-weight is regained at about the 16th year.¹ The marked immediate postnatal depression of growth shown by the adrenals is not shared by other endocrine glands, e.g., the thymus, pineal, pituitary, thyroid, ovaries and testes. According to the evidence reviewed by Parks (15), the adrenal glands of the neonate are composed almost entirely of cortical tissue, and their comparatively large size derives from the marked development during the fetal period of what is termed the fetal cortex, which is a special temporary inner cortical zone. Following birth this zone degenerates and gives place to the *zona reticularis*. The essential structural features of the adult adrenals are found by the end of the second month. It is further known that while the adrenal cortex derives from the celomic mesoderm, the medulla is ectodermal in origin.

It appears that there is some basis for supposing a positive relationship between the size of a gland and the amount of its secretory activity during the course of ontogenetic development. The testes and ovaries, for example, in-

crease rapidly in size with the onset of sexual maturity. In the case of the adrenals, however, we find that the postnatal loss of weight can be attributed in part to the degeneration of the fetal cortex. If our evidence is correct, the fetal cortex gives place to the *zona reticularis* by the end of the second mo. Yet loss of weight of the glands continues to the second year after birth. Our basic question is whether the comparatively small size of the adrenal glands in infancy and childhood is actually associated with relative lacks of endocrine products assumed to be important in emotion. In undertaking an answer to this question it is necessary to consider first the evidence showing that both the adrenal medulla and the adrenal cortex respond to stress.

Cannon's (5) conclusions regarding the role of the medullary hormone, adrenin, in emotion are well known to psychologists. His extensive researches appear to have led most authorities to accept his fundamental conclusions regarding the dynamogenic effects of adrenin. His conclusions, however, have been challenged. According to Rogoff (17), many of the reactions attributed to adrenin can be obtained in animals in whom adrenin secretion has been abolished. Furthermore, such animals show a persistence of emotional states. Arnold (1) has reviewed evidence indicating that adrenin depresses rather than raises somatic activity, and has proposed that fear, which is enervating rather than energizing, is the emotion involving predominantly sympathetic excitation. Anger, on the other hand, which is associated with general activity is supposed by Arnold to involve strong parasympathetic excitation. An adequate discussion of the objections raised by Rogoff and Arnold to Cannon's assumptions cannot be undertaken here. It would be necessary for us to consider in detail answers to questions of the fol-

¹ Scammon's portrayal of the growth of the adrenal glands differs somewhat from that of Jackson (13). Both indicate the same general depression extending to the 2nd year. Jackson's curve, however, is relatively flat between the ages of two and six, and indicates that the birth-weight is regained at about the 11th year. These differences are not crucial from the standpoint of this discussion.

lowing type: To what extent is the organism capable of showing *adrenin-like* reactions in the absence of adrenin? May we say that the reduction of overt activity sometimes incident to fear actually involves depression of somatic activity? To what extent is it possible to estimate the amount of adrenin normally present in the blood without the consequences of stress incident to this determination? In short, we believe that though we are not committed to the full acceptance of Cannon's emergency theory of emotion, his major conclusions regarding the somatic effects of adrenin in emotion remain tenable. We note, further, that the liberation of adrenin incident to stress is relatively rapid. Selye (21, p. 127) states this as follows: "Hyperadrenalinemia is characteristic of the first stage of the alarm reaction, and accounts for many of the manifestations seen in this initial phase of a general adaptation syndrome." As far as adrenin is concerned, it would appear that the liberation of adrenin into the blood stream enhances the vigor of both transitory emotional states, and the initial phases of prolonged emotional states.

Our survey of recent hormone research on the adrenals indicates that the main interest has shifted from the adrenal medulla to the adrenal cortex. Pincus (16) has assembled findings from a group of studies of the role of the adrenal cortex in stress of human subjects. These studies indicate that hypersecretion of cortical hormones occurs as a consequence of experimentally imposed stresses such as those induced by pursuit meter operation and airplane flying, and by stresses (emotionally toned) encountered in daily life. The amount of adrenal cortex activity is inferred on the basis of urinary 17-ketosteroid output (most of which may be traced to the adrenal cortex steroid), and lymphocytopenia. The adrenal

cortex, unlike the adrenal medulla, is not subject to direct neural regulation. It appears established, however, that the anterior pituitary exerts regulatory control of the adrenal cortex through the release of an adrenocorticotrophic hormone (ACT), and it is this hormone that brings about the release of steroid hormones. Thus as Selye (21) has indicated, essentially the same agents influence the production of both adrenin and corticoids, but the former is liberated more rapidly than the latter. Over 25 different steroid compounds have been isolated from concentrates of the adrenal cortex. The roles of these compounds in bodily economy are complex, and the evidence does not appear to justify definitive statements regarding these roles. For the purposes of the present discussion, it would seem appropriate to assume the validity of relatively gross generalizations. We may appropriately cite certain concluding statements made by Sayers and Sayers in their discussion of the pituitary-adrenal system:

"The response of the peripheral tissues to stressful conditions is characterized by an increased utilization or destruction of cortical hormone(s). The functional change in tissue cells produced by the great variety of nonspecific stresses which initiates this increased utilization is as yet unknown. . . . It has been pointed out that there appears to be an association between cortical hormone utilization and the general level of metabolic activity of cells" (19, pp. 104-105).

The requirement of tissue cells for cortical hormones during stress appears to be great. Large amounts of these hormones would therefore seem to be necessary for the maintenance of prolonged emotional states.

There remains the important question of whether the comparatively small size of the adrenal glands in early life is actually associated with relative defi-

ciencies in hormone output. So far we have been unable to find evidence bearing on the relative magnitude of adrenin secretion in early life. Unlike the medulla, the adrenal cortex is necessary for the preservation of life. As far as the adrenal cortices are concerned, the problem is that of whether some of them are relatively deficient in early life. We have already noted the elevation in urinary 17-ketosteroid output which is a consequence of stress. Talbot and Sobel (22) state that 17-ketosteroids are not produced in measurable amounts until the ninth or tenth year of life. While we would submit that the cortical hormones may have different metabolic end-products in children, Talbot and Sobel conclude that the 17-ketosteroid precursors are of minor importance for the maintenance of the rate and duration of growth in normal childhood. From our point of view it seems plausible to suppose that while the apparent deficiency in adrenal cortices does not prevent the early occurrence of emotional behavior, it may influence the nature of early emotional states.

From the above discussion we emerge with the following conclusions regarding the two original assumptions. Evidence on the development of the nervous system and developmental studies of the EEG support the assumption that the infant is a precorticate organism and cortical influence, especially that of the frontal lobes, develops slowly. The consequences of this delayed development are reflected in the characteristics of early emotional behavior. Support for the assumed roles of the higher centers of the nervous system is found in the consequences of the experimental decortication of animals and pre-frontal lobotomies. The assumption of a relative lack of endocrine products having significant roles in emotion is suggested by available evidence but not proven.

The adrenal glands, unlike other endocrines, show a peculiar early retardation in their gross development. That this retardation is actually associated with relative hormone deficiencies cannot at present be established, though the infant appears to show no measurable amounts of the 17-ketosteroids which derive from precursors in the adrenal cortex. The significance of this deficiency derives from the established correlation between stress and the output of urinary 17-ketosteroids. This treatment of the ontogenesis of emotional behavior obviously emphasizes the role of maturation. Insofar as learning is necessarily involved in the patterning of behavior, we would simply point out that the learning process should be analyzed within the limits imposed by maturation.

Our discussion is obviously oversimplified, and we have avoided a discussion of the neural mechanisms that constitute emotion. Our theoretical assumptions have the merit, we believe, of indicating areas of needed research. Further studies should be made of emotional responses in childhood, especially of the incidence, magnitude and duration of internal changes. These studies should include chemical assays of hormone output. It appears likely that other glands than the adrenals, especially the pituitary, will be found to provide useful evidence on the physiology of emotion. Further studies should also be made of the emotional responses of decorticate animals, cases of extreme mental deficiency, and persons who have undergone frontal lobotomy. Perhaps after such research is done the concept of emotion will become definable, and it will cease to be such a "whale among the psychological fishes."

REFERENCES

1. ARNOLD, M. B. Physiological differentiation of emotional states. *PSYCHOL. REV.*, 1945, 52, 35-48.

2. BARD, P. Emotion: I. The neuro-humoral basis of emotional reactions. In Murchison, C. (ed.), *Handbook of general experimental psychology*. Worcester, Mass.: Clark University Press, 1934. Pp. 264-311.
3. —, & MOUNTCASTLE, V. B. Some fore-brain mechanisms involved in the expression of rage with special reference to the suppression of angry behavior. *Res. Publ. Ass. nerv. ment. Dis.*, 1947, 27, 362-404. (Quoted by C. T. Morgan & E. Stellar in *Physiological psychology*, New York: McGraw-Hill, 1950, pp. 353-354.)
4. BRIDGES, K. M. B. Emotional development in early infancy. *Child Developm.*, 1932, 3, 324-341.
5. CANNON, W. B. *Bodily changes in pain, hunger, fear and rage*. New York: Appleton, 1929.
6. FLECHSIG, P. *Anatomie des menschlichen Gehirns und Rückenmarks auf myelogenetischer Grundlage*. Leipzig: Thieme, 1920. (Quoted by W. Freeman and J. W. Watts, *Psychosurgery*, 1942, p. 24.)
7. FORD, F. R. *Diseases of the nervous system in infancy, childhood and adolescence*. Springfield, Ill.: Thomas, 1937.
8. FREEMAN, W., & WATTS, J. W. *Psychosurgery*. Springfield, Ill.: Thomas, 1942.
9. GIBBS, F. A., & GIBBS, E. L. *Atlas of electroencephalography*. Cambridge, Mass.: Cummings, 1941.
10. HARLOW, H. F., & STAGNER, R. Psychology of feelings and emotions: I. Theory of feelings. *PSYCHOL. REV.*, 1932, 39, 570-589.
11. HEBB, D. O. Animal and physiological psychology. In Stone, C. P. (ed.), *Annual Review of Psychology*. Stanford, Cal.: Annual Reviews, 1950.
12. —. *Organization of behavior*. New York: Wiley, 1949.
13. JACKSON, C. M. Some aspects of form and growth. In Robbins, W. J., Brody, S., Hogan, A. G., Jackson, C. M., and Green, C. W., *Growth*, New Haven, Conn.: Yale University Press, 1928. Pp. 109-140.
14. JERSILD, A. T. Emotional development. In Carmichael, L. (ed.), *Manual of Child Psychology*. New York: Wiley, 1946. Pp. 752-790.
15. PARKS, A. S. Adrenal-gonad relationship. *Physiol. Rev.*, 1945, 25, 203-254.
16. PINCUS, G. Studies of the role of the adrenal cortex in the stress of human subjects. In Pincus, G. (ed.), *Recent progress in hormone research, Vol. I*. New York: Academic Press, 1947. Pp. 123-145.
17. ROGOFF, J. M. A critique on the theory of emergency function of the adrenal glands: implications for psychology. *J. gen. Psychol.*, 1945, 32, 249-268.
18. RUSSELL, W. R. Functions of the frontal lobes. *Lancet*, 1948, 254, 356-360.
19. SAYERS, G., & SAYERS, M. A. The pituitary-adrenal system. In Pincus, G. (ed.), *Recent progress in hormone research, Vol. II*. New York: Academic Press, 1948. Pp. 81-115.
20. SCAMMON, R. E. The measurement of the body in childhood. In Harris, J. A., Jackson, C. M., Paterson, D. G., and Scammon, R. E., *The measurement of man*. Minneapolis, Minn.: University of Minnesota Press, 1930. Pp. 171-215.
21. SELYE, H. *Textbook of endocrinology*. Montreal: University of Montreal, 1947.
22. TALBOT, N. B., & SOBEL, E. H. Certain factors which influence the rate of growth and the duration of growth of children. In Pincus, G. (ed.), *Recent progress in hormone research, Vol. I*. New York: Academic Press, 1947. Pp. 355-369.

[MS. received November 15, 1950]

PRIMARY STIMULUS GENERALIZATION: A NEUROPHYSIOLOGICAL VIEW

BY JOSEPH WOLPE

University of the Witwatersrand, South Africa

In two previous articles some fundamental facts of stimulus patterning (13) and of learning (14) were interpreted in terms of current neurophysiological knowledge. The present paper deals in similar fashion with primary stimulus generalization.

It is a commonplace observation that when an organism has been conditioned to make a response to a clearly defined stimulus, it tends to make a similar response to a second stimulus similar to the one conditioned. This is what Hull (5, pp. 184-187) has called *primary stimulus generalization*, for it is a direct result of "the partial physical identity of the stimulation compounds." The physical similarity between two stimulus compounds may be discernible in any of a number of "dimensions"—for example, frequency of vibrations, shape, position, or identity of parts.

The view to be put forward here is that primary stimulus generalization can be accounted for by the fact that similar stimuli excite a certain number of afferent neurones in common, this number decreasing as the two stimuli diverge from each other in any dimension. It is well known that the greater the resemblance of the generalized stimulus to the conditioned stimulus, the stronger the response that the generalized stimulus can evoke. Two examples of this are given below, and in each case there is evidence of a parallelism between the degree of similarity of the stimuli and the number of neurones they excite in common.

In the case of tactile stimuli Pavlov (7, p. 113), Anrep (2), and Bass and Hull (3) have found that if a response

is conditioned to a stimulus to a given spot on the skin, application of the stimulus to other spots will also elicit the response, but with a strength that decreases with increasing distance from the spot where conditioning was established. This inverse relationship fits in neatly with the fact first noted by Sherrington (8, pp. 126-129) that nerves supplying closely adjacent cutaneous spots give off impulses to more spinal afferent neurones in common than is the case when the spots are more widely spaced.

With regard to auditory stimuli, Pavlov (7, pp. 118-127), Hovland (4), and Humphreys (6) have reported that the responses to a stimulus differing in pitch from a conditioned auditory stimulus diminish in magnitude as the difference in pitch becomes wider. (Humphreys, it is true, having established a conditioned galvanic response to a tone of 1967 cycles, found a significantly greater generalized response to 984 cycles than to 1000 cycles. This, however, is easily understood, since 984 cycles, being an octave below 1967 cycles, would produce overtones of the latter frequency.) That such auditory generalization also depends on the number of neurones that are activated in common by different frequencies is indicated by the findings of Stevens, Davis and Lurie (12) that the lower the frequency of a tone acting on the ear, the nearer to the apex is the responding part of the organ of Corti, and the greater the similarity between the tones, the greater the overlap between the respective responding parts. It has further been demonstrated that adjacent parts of the organ of

Corti are connected with adjacent areas of the primary acoustic nucleus and the medial geniculate body (1) and with adjacent areas of the cerebral cortex (15).

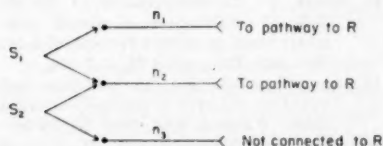


FIG. 1. Response R has been conditioned to stimulus S_1 which activates neurones n_1 and n_2 . S_2 , also activating n_3 , also elicits R , though less strongly than S_1 .

In the above two examples there is good support for the suggestion that primary stimulus generalization is accounted for by similar stimuli exciting a certain number of neurones in common. Of course, experiments which would demonstrate this more directly are very desirable. Also desirable is information which would make possible correlations involving other dimensions of difference and other sensory modalities. Nevertheless, we already seem to have fairly satisfactory grounds for the following general hypothesis:

Primary stimulus generalization occurs when a stimulus S_2 , not previously conditioned to a response, evokes that response in some measure by activating some of the afferent neurones that are also activated by a stimulus S_1 that has been conditioned to the response; and the strength of the response to S_2 will depend upon the number of neurones whose activation is common to both stimuli (see figure).

A Corollary: A corollary of this hypothesis is that common pathways are also responsible for the generalization of extinction effects. It may be noted that this would flatly contradict Spence's expectation (9, 10, 11) that if the generalized response to a stimulus close to the conditioned stimulus is extinguished,

the generalized response to a stimulus close to the other side of the conditioned stimulus in the same dimension will be found to be stronger than the response to the conditioned stimulus itself. For instance, if a response is conditioned to a tone of 1000 cycles, and then extinction of that response to 950 cycles is accomplished, afterwards, according to Spence, the response will be stronger to 1050 cycles than to 1000 cycles. According to the present neurophysiological hypothesis this is impossible, for the extinction of the response to a stimulus acting through generalization means the inactivation of some of the "marginal" pathways reinforced to the response. The original conditioned stimulus would continue to act through *all remaining pathways*. Any stimulus whose effect depended on generalization would have at its disposal fewer pathways to the response, and should therefore evoke a weaker response. A crucial experiment on this matter does not yet appear to have been done. Spence's own experiments are primarily determinations of the relative extent to which a given response is due to the absolute effects of a stimulus, and to that stimulus as part of a pattern.

REFERENCES

1. ADES, H. W., METTLER, F. A., & CULLER, E. A. Effect of lesions in the medial geniculate bodies upon hearing in the cat. *Amer. J. Physiol.*, 1939, 125, 15-23.
2. ANREP, G. V. The irradiation of conditioned reflexes. *Proc. Roy. Soc. London*, 1923, 94B, 404-426.
3. BASS, M. J., & HULL, C. L. The irradiation of a tactile conditioned reflex in man. *J. comp. Psychol.*, 1934, 17, 47-65.
4. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, 17, 125-148.
5. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.

6. HUMPHREYS, L. G. Generalization as a function of method of reinforcement. *J. exp. Psychol.*, 1939, **25**, 361-372.
7. PAVLOV, I. P. *Conditioned reflexes* (trans. by G. V. Anrep). London: Oxford University Press, 1927.
8. SHERRINGTON, C. S. *The integrative action of the central nervous system*. Cambridge: Cambridge University Press, 1947.
9. SPENCE, K. W. The differential response in animals to varying stimuli within a single dimension. *Psychol. Rev.*, 1937, **44**, 430-444.
10. —. Failure of transposition in size discrimination of chimpanzees. *Amer. J. Psychol.*, 1941, **54**, 223-229.
11. —. The basis of solution by chimpanzees of the intermediate size problem. *J. exp. Psychol.*, 1942, **31**, 257-271.
12. STEVENS, S. S., DAVIS, H., & LURIE, M. H. The localization of pitch perception on the basilar membrane. *J. Gen. Psychol.*, 1935, **13**, 297-315.
13. WOLPE, J. An interpretation of the effects of combinations of stimuli (patterns) based on current neurophysiology. *Psychol. Rev.*, 1949, **56**, 277-283.
14. —. Need-reduction, drive-reduction and reinforcement: a neurophysiological view. *Psychol. Rev.*, 1950, **57**, 19-26.
15. WOOLSEY, C. N., & WALZE, E. M. Topical projection of nerve fibres from local regions of the cochlea to the cerebral cortex of the cat. *Bull. Johns Hopkins Hospital*, 1942, **31**, 315-344.

[MS. received November 15, 1950]

A PREFACE TO A PSYCHOLOGICAL ANALYSIS OF THE SELF¹

BY THEODORE R. SARBIN

University of California

The purpose of this preface is fourfold: *first*, to suggest that much of the undiagnosed confusion in the theory and practice of psychology is due to the multiplicity of meanings attached to the words self and ego; *second*, to sketch briefly a theory of self-development (which I have called epistemogenesis because it emphasizes cognitive factors); *third*, to illuminate the subject-object problem with the aid of this theoretical approach; and *fourth*, to formulate an illustrative hypothesis the testing of which would give at least initial validity to the theory.

Fifty years ago a psychological theorist would have found it hard going to write a psychology that did not deal in one way or another with the self. Such keen thinkers as James (15), Titchener (33), Hall (14), McDougall (21), Baldwin (3), and many others, posited a self or ego as a conception without which psychological theory just wouldn't make sense. Although they all agreed that such a concept was necessary, they differed amongst themselves as to the nature of the self, its development, and its function in various psychological processes. For example, Titchener declared

simply that the "self is the sum total of conscious processes which run their course under conditions laid down by bodily tendencies" (33), while William James in more complicated manner described the self or "me as an empirical aggregate of things objectively known. The [Ego or] I which knows them . . . is a *Thought*, at each moment different from that of the last moment . . ." (15).

Except for certain European psychologists, such as Stern (31) and Claparède (8), the psychoanalytic theorists, notably Freud (11) and Jung (17), and such sociologically oriented scholars as Cooley (9) and Mead (22), the self all but disappeared from learned psychology with the rise of behaviorism. The emphasis on the exclusively objective approach to psychological phenomena considered the self (with its soulful antecedents) not essential in formulating psychological theory.

The recrudescence of the self began with Allport's book on personality (1) and since that time the ego has become more and more popular in psychological discussions.² In fact, in 1943, Allport felt called upon to comment on the multiplicity of meanings attached to the term (2). In describing eight general meanings that were current, he suggested that one all-inclusive theory of the ego might account for the multiplicity of functions and characteristics attributed to it. Although beginnings have been made (7), the all-inclusive theory has yet to be advanced.

To begin this analysis, some of the

¹ No. 3 in a series of papers entitled *Contributions to Role-taking Theory*.

A somewhat abbreviated version of this preface was read at the 1950 meetings of the Western Psychological Association. For many stimulating comments I wish to thank the members of an informal graduate seminar: Messrs. N. Adler, D. R. Brown, N. S. Greenfield, M. Hyman, W. F. McCormack, and R. Taft. Mr. Greenfield should be credited with first suggesting the term "epistemogenesis." I am also grateful to Professor T. M. Newcomb for many valuable criticisms.

² Self and ego are here used synonymously.

main conceptions of the ego were reviewed and at least a dozen variants were discovered. Some of these are: the physical self, the material self, the introjecting ego, the empirical self, the projective self, the pure ego, the transcendental ego, the social self, the ethical self, the inferred self, and so on. In the contexts in which these expressions were used, each had some degree of validity, yet they were obviously not the same things. For example, one would have to stretch his credulity to equate the somatic self of the neonate with the ethical self of a university professor struggling with the problem of signing or not signing a controversial document. It is unnecessary to labor the point further; a large number of terms are in use that appear to have a common signification, but communication on the subject is so lacking in unity, that confusion is often a result.

OUTLINE OF THE EPISTEMOGENIC THEORY

With the aforementioned multiple referents in the background, the formulation of a theory of the origins and development of the self is herewith essayed, being guided by the following informal postulates:

1. The human animal can regard itself as an object in the same way as it regards objects in the external world. Put in other terms, the interbehavioral field of the human can include perceptions and cognitions referable to objects in the external world, and perceptions and cognitions referable to his own body, to his own beliefs, his own statuses, and so on.

2. Behavior is organized around cognitive structures, the result of responses of the organism to stimulus-objects and residual stimuli. The self is one such cognitive structure or inference. Like all cognitive structures it is organized around substructures, here called em-

pirical selves. These substructures are interrelated through some learning mechanism.

3. The self is empirically-derived, not transcendental; it is the resultant of experience, i.e., interaction with body-parts, things, persons, images, and so on.

4. The properties of these substructures, at any given moment are determined by the total interbehavioral field of which the substructures are a part. Thus, any of the empirical selves may occupy the focus of the interbehavioral field at any given time.

5. The self (in common with other cognitive structures) is subject to continual and progressive change, usually in the direction from low-order inferences about simple perceptions to higher-order inferences about complex cognitions.

6. Organic maturation and reinforcement of selected responses contribute to changes in the empirical selves (substructures) and concurrently to changes in the total self-structure.

7. Change in cognitive structure is a function of (at least) two properties: (a) resistance to de-differentiation and (b) the breadth of the substructures. Resistance to de-differentiation may be defined in terms of the strength of the boundaries of a structure which, in turn, is related to overlearning. Breadth may be considered the dimension that determines how many structurally-dissimilar perceptions may be included in the same substructure. This view is similar to Frenkel-Brunswick's notion of ambiguity-tolerance (10).

With these guideposts set down, an initial problem is approached. Is the self as a cognitive structure to be considered an enduring structure, a momentary structure, or both? Calkins' definition of the self includes the notion that the self is an enduring structure: "The characters of the experienced self . . . are, first, its persistence or self-

identity, second, its individuality or uniqueness, third, . . . it is fundamental . . . to its experiences, and finally, . . . it is related to its environment" (6, pp. 493-494). The self-structure as a cross-section in time can be inferred from G. Stanley Hall's statement: "The earliest parts of the physical self to attract attention are the hands and fingers. . . . Children of four and five months are described as attentively feeling of one hand with the other, each at the same time feeling and being felt, each subject and object to the other, and thus detaching them from the world of external things and labeling them with a mark which will enable the soul later to incorporate them into the plexus which forms the somatic ego" (14, pp. 351-352). Statements of this kind, which influenced the development of the postulates, stimulated the construction of a graphic model which allowed for the simultaneous presentation of the self as a longitudinal structure and as a cross-sectional structure (see Fig. 1).

The model is divided into sections so that the reader may more easily follow the argument. Figure 2 overlaps Fig. 1, Fig. 3 overlaps Fig. 2, etc. Figure 6 is a schematized outline of the part of the theory which is elaborated in this preface. Items of behavior are introduced along the continuum to illustrate how

certain empirical data fit into the differentiated substructures. The age-references are approximate.

In this analysis five substructures are developed; these substructures, or empirical selves, may be regarded as stages of refinement in discrimination of stimuli, or as the development of differential reference-schemata. The second reference-schema, the receptor-effector self, is an emergent which is formed after some consolidation (fixing of boundaries) has taken place in the first reference-schema. Similarly, the third empirical self appears after the second is consolidated, etc. Overlearning (fixation) at any level may produce reference-schema whose boundaries are so rigid that all incoming stimuli are cognized in terms of those schema so far developed, at the same time inhibiting or delaying the formation of new cognitive structures.³

I. At the apex of the cone is the neonate. Because the stimulus field is probably made up exclusively of unorganized somatic sensations (see Fig. 1) any "inferences" or cognitions made by the infant would be unstable and of the most primitive and rudimentary variety.

³Only the main outlines of the cognitive structures are sketched. In the larger work to be published later, the several substructures are described more systematically and in greater detail.

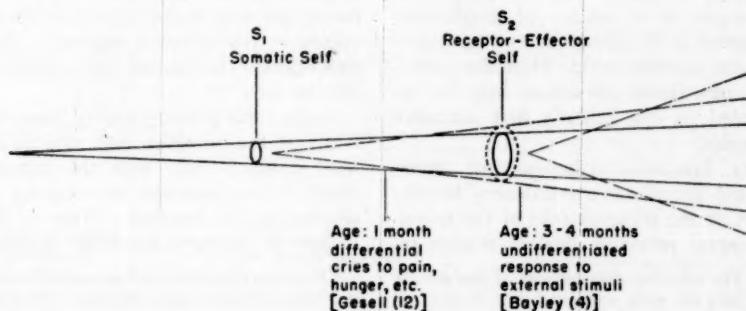


FIG. 1.

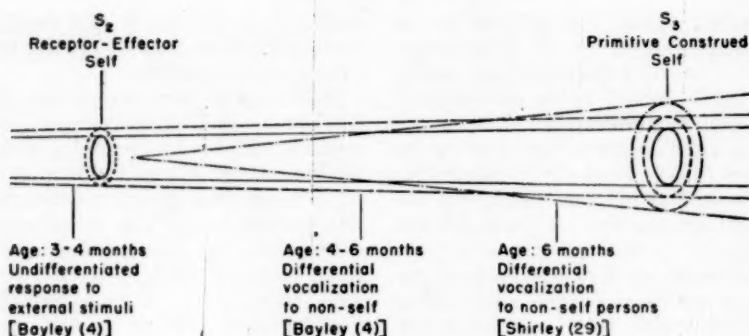


FIG. 2.

Available data indicate that CR's—which may be regarded as the most rudimentary type of inferential behavior—while possible, are difficult to establish during the first month and are easily extinguished (23). The first perceptual organization of the infant is probably in relation to the modes of restoring homeostatic balance at the physiological level. The *somatic self* becomes organized around responses initiated by stimuli to the somesthetic senses, kinesthesia, proprioception, and the cutaneous senses. Most of the infant's motility, for example, is accounted for by internal stimuli (25). That there is initially no differentiation of self from non-self may be inferred from casual observation of the neonate during the feeding period. What he grasps, for example, is a matter of indifference whether it be a body part or an object in the outside world. Thus the matrix of somesthetic sensations may be regarded as the infant's first reference schema.⁴

II. Tensions arising out of uncorrected homeostatic imbalances become part of the stimulus field of the infant. Receptor processes lead to tension re-

duction—the paradigm of these processes is the incorporation of nutriment. With the maturation of the organism and with its increased motility, other kinds of stimuli are mediated via the distance and other receptors. Concurrently, tensions introduced into the infant's stimulus field by the accumulation of waste products are reduced by motor activity of the skeletal muscles and sphincters. In the infant's reference-schemata there is probably no distinction between *objects* which are instrumental in tension-reduction and *persons* who are instrumental in tension-reduction. We might infer that to the infant certain perceptions are associated with tension-reduction. At this point, of course, tension-reduction is a mass-action or global affair—probably all the perceptual and motor apparatus is involved in any behavior segment. This substructure is labeled the *receptor-effector self*.⁵

Around this primary matrix, then, the infant at the primitive deductive (CR) level interacts only with the non-self which is instrumental in reducing or eliminating its tensions. This is the *Anlage* of the next-described cognitive

⁴ The cognitive development of the somatic self into the more complex "body image" will be described in the fuller treatment of the theory.

⁵ The names attached to these cognitive substructures are admittedly awkward. In seminar discussions the substitution of symbols S_1 , S_2 , S_3 , S_4 , and S_5 has been helpful.

structure which begins to differentiate between non-self objects and non-self persons (see Fig. 2).

III. The next level of development is labeled the *primitive construed self* (see Figs. 2 and 3). At the preceding level the cognitive structure is not sufficiently differentiated to perceive differences between *objects* which are instrumental in tension-reduction and *persons* who are instrumental in tension-reduction. Now, with increased physiological maturation and with the effects of reinforcement of certain responses selected by others, the cognitive organization differentiates *between objects and persons*. Apparent to the observer at this level is the differential reactivity of the infant. For example, he vocalizes pleasure, eagerness, and displeasure (4, 20). He is active, he attempts to touch, he discriminates tone and expression, he distinguishes persons (4, 20). The important distinction between this level of selfhood and the previous levels is that here the cognitive structure embraces perceptions of persons who are selective, whimsical, and dynamic. The stimulus properties of persons are ever-changing affairs, as contrasted to the stimulus properties of things. Baldwin's analysis is apposite

here. "... this is the child's very first step toward a sense of the qualities which distinguish persons. The sense of uncertainty grows stronger and stronger in its dealings with persons. A person stands for a group of experiences quite unstable in its prophetic as it is in its historical meaning. This we may, ... in assuming it to be the first in order of development, call the 'projective stage' in the growth of the child's consciousness" (3, p. 7). Thus, the child's reference-schemata must now function to select the more perduring stimulus qualities from the more evanescent. Because of physiological maturation and also because of the acquisition and the consolidation of a schema for perceiving many objects, there now emerges a new cognitive substructure, a new reference-schema, a new self. This emergence is facilitated by non-self persons who, unlike objects which serve mainly on the stimulus side, interact with the child in *both* the stimulus and the response aspects. There is still relatively little binding of tension—no appreciable delay between sensory and motor aspects: Perception of stimuli and motor discharge are relatively continuous proc-

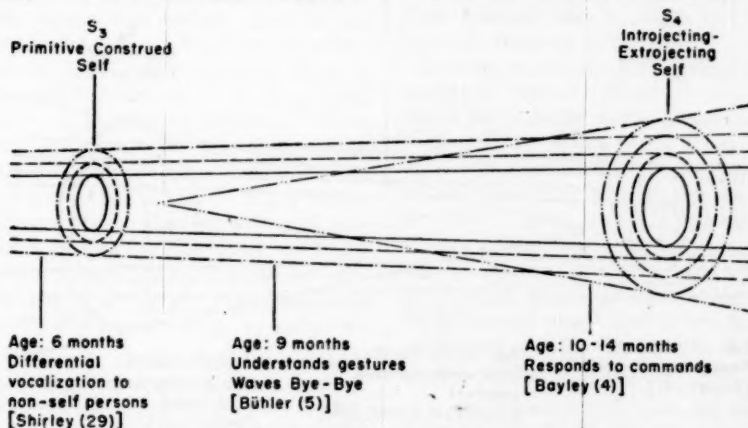


FIG. 3.

esses.⁶ Available data indicate that delay, followed by responses indicating memory, is not more than one minute in the 10-11-month-old child (23). At 6-8 months, time-binding is probably much less. Perception is still of the physiognomic kind, literal perception depending upon greater differentiation in cognitive structures (35).

At this level, then, we can infer that the child has three main substructures or reference-schemata: the somatic self, the receptor-effector self, and the primitive construed self. These are all empirical—all derived from interaction with things, people, and perhaps simple signs. With these cognitive structures,

⁶ These tensions of course stimulate the mother or nurse to aid in tension-reduction. The cries, restlessness, etc., exhibited by the baby are treated as a signal by the adult. Since the ministrations of the adults which follow the child's "signals" have some regularity and organization, the infant might formulate—in a primitive way—a deductive inference which has a magical denotation such as the "superstitions" developed by Skinner's pigeons (30). This parallel between the "superstitions" of these pigeons and the magical thinking of adult persons whose self-structures are dominated by this empirical self is utilized in the illustrative hypothesis below.

the infant interacts with non-self objects and non-self persons both of which are instrumental in reducing tensions.⁷

IV. Cognition becomes more varied, rich, and complex with the growth of non-verbal and verbal language structures. At the next level of the self, the substructures so far described are central to a new emergent—the *introjecting-extrojecting self* (see Figs. 3 and 4). The distinguishing feature here is the use of crude language devices in isolating, differentiating, and strengthening the cognitive structure *which the child communicates to others*. At the same time, development of motor skills, increased motility, postural changes, etc.,

⁷ T. M. Newcomb (personal communication) suggests calling these three aspects of selfhood "pre-self" differentiations, reserving the term "self" for those later differentiations which depend on the more refined use of symbols. While this emendation would be more continuous with the traditional sociological usage of the term, it might be interpreted as signifying a discontinuous process. It should be emphasized that although the earlier cognitive structures are rooted in the more somatic aspects of behavior and are derived from interaction with signs, they serve as the *Anlage* for the later cognitive structures which involve interaction with symbols.

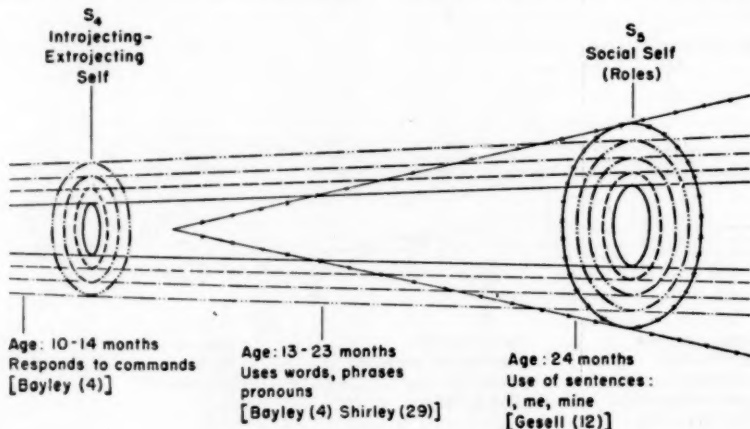


FIG. 4.

increase the number and range of potential non-self objects and persons that can be responded to and integrated into the currently-forming cognitive structure. In addition, the child can now differentiate words, imitate words, cooperate in play, etc. (4, 20).

In the interbehavioral field are different persons whose responses to the child are variable. It is probably at this level that Baldwin's "dialectic of personal growth" (3) has its greatest applicability. The reactions to the multiple and shifting stimulus properties (acts) of others are introjected or assimilated (or they become "subjective" according to Baldwin) if congruent with the reference-schemata so far developed. Extrojection (Baldwin's "ejective consciousness") probably emerges simultaneously. The child's own responses to the differential acts of other persons are organized into this substructure. By a primitive analogical process which develops with the increase in language structures the acts of others are perceived as the same as the child's own acts.*

Introjecting and extrojecting responses become more differentiated as a result (1) of increased locomotor and manipulative skills and (2) of the selection and reinforcement of certain types of responses by others. The most important developmental feature at this stage is the refinement in the use of the gesture (22). First, motoric gestures, "whole-body language" (19), and, later, vocal gestures become the instruments by which the child organizes this new reference-schema. The importance of language in concept formation needs no comment here. With language symbols, the acts of others can more readily be conceptualized as similar to one's own

acts (projection) or one's acts can more readily be conceptualized as similar to the acts of others (identification).

The child has not learned to say I, although he does use his own name as others use it; he employs other gestural equivalents which are imitations of others' behavior in referring to him (12). Here we have the *Anlage* of the social self.

V. With the development of the first person personal pronoun, I, the child attains a more refined concept than formerly, at the same time acquiring a conventional symbol for self-reference (see Fig. 4). With I, mine, and me, he can categorize perceptions in a more sharply differentiated way than before. The *social self* thus emerges—a new cognitive substructure. The child has a reference-schema with which he can organize perceptions and cognitions of the organized behaviors of other persons. Now he can differentiate not only *discrete acts* of persons, but *organized acts* or *roles*.

The conventional role of mother and the conventional role of father, e.g., are easily differentiated. Differential roles are perceived and integrated into the cognitive structure. Roles are seen as differentiations of this substructure because different roles in others call out different responses in the child.

Because perception of a stimulus and enactment (*motor discharge*) are so closely tied together in the development of the self up to this point, *acting* the role of the other implies the concurrent or just prior *perception* of the role of the other. Later, *delay* is introduced between role-perception and role-enactment.⁹ Acting the role of the other often means attaching adult values to perceptions of others and indirectly to perceptions of self. The further development of the self-structure into the

* Weiss has written a technical but readable description of these processes from the psychoanalytic viewpoint. Compare his concepts of introjection, extrojection and objectification with this portion of our analysis (34).

⁹ A second paper will deal with this aspect of the self with special reference to motivational analysis.

"generalized-other" (22) could be represented by the emergence of another substructure. Because of the limitations imposed by the prefatory nature of this essay, a discussion of further development of the self will be deferred.

Like other models, this one has its deficiencies. Figure 5 shows the somatic self as the core of the self-structure with the other selves concentric to it. While the somatic self may be the central feature of the self-structure in early childhood, it probably becomes subordinated to other aspects of selfhood in later life. Therefore, a transformation is necessary. The various selves are best regarded as semi-independent substructures which later become overlaid with substructures of the social self—the roles (see Fig. 5).

This is not the terminus of the epistemogenic theory. In this essay we are content to point out the directions that must be taken. A more extended treatment of the theory will be made available later. In fact, the model represents only the beginning of the self-organization which becomes tremendously hypertrophied and complicated by further differentiations in the cognitive structure due to the role-taking process.

Additional reference-schema are formed as a result of the interaction of the self with others. From this point on the epistemogenic theory is carried forward with the aid of role-taking concepts (22, 26, 27, 28). In a word, roles which are perceived as congruent with one's current self-organization are capable of enactment; roles which are incongruent with the structure of the self are distorted in enactment or delayed or rejected.

Subject or Object? Current discussions of the self often distinguish between the knower and the known or between the subjective and objective aspects of the ego or self. William James (15) and Mead (22) both treated these as the "I" and the "me." This distinction is usually equated with private and public aspects of selfhood. That this distinction is not valid is seen from this example: A person's fantasies, while private, and supposedly subjective, may be treated as an object. The person, by virtue of his cognitive organization, may identify and evaluate his fantasy reactions in the same way that he identifies and evaluates external objects. Therefore, to consider the self as made up of subjective and objec-

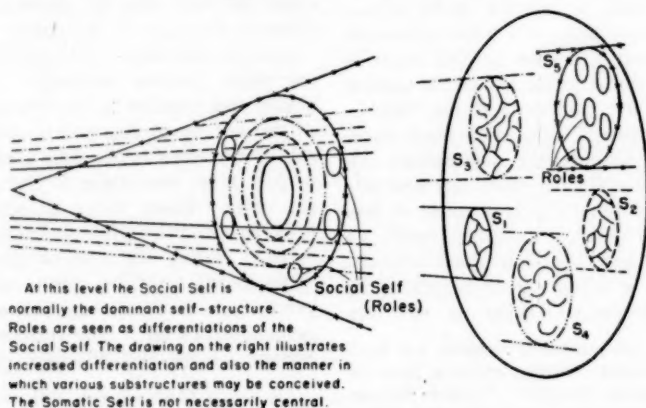


FIG. 5.

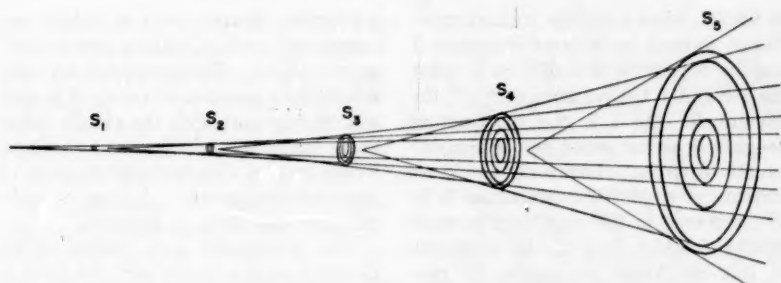


FIG. 6. Schematized outline of epistemogenic theory.

tive aspects is an over-simplification. A dualistic set of postulates is implied in such a formula. Most contemporary discussions, notably Symonds (32), are able to give a clear and straightforward presentation of the development of the self as object, but the self as subject, as knower, doer, thinker, etc. is considered a given or is not considered at all.

In the epistemogenic theory the I and the me are both handled in terms of a consistent set of postulates without having recourse to the old dualistic tradition. The self as I is seen as an *inference*, a high-order inference or cognitive structure, which develops as a result of and consists of lower-order inferences (or reference-schemata) which are here called empirical selves. To support the notion of its being a higher-order inference, that is to say, one which develops out of the matrix of lower-order inferences, we have only to mention again the studies of language development in children. The consistent use of the concept "I" occurs relatively late—about two years (12).

This theory takes its point of departure from James (15) and Pillsbury (24). Both of these writers saw the problem in terms of empirical selves as the "objects of awareness," and the I or pure ego as "interpretations" (Pillsbury) or "Thought" (James). In terms of the present theory, when a person

uses the term "I" the referent is a cognitive structure, an inference, the internal organization of which is characterized by substructures with varying properties of strength and breadth.

Thus, a patient with strong internal pressures for constant massage (somes-
thetic stimulation) might have a concept of self in which the earliest empirical self is the most developed substructure—the somatic self. His use of the pronoun "I" then is referable to a cognitive structure (a cross-section in the immediate present of the entire interbehavioral field) in which the somatic self is dominant. Similarly, when another person says, "I am always craving food," the I may have as its referent a conceptual organization in which the empirically-derived substructure known as the receptor-effector self is in focus. Or, when a person says "I can't understand why people are always losing patience with me," we might suggest that the referent for the I is a cognitive structure where the primitive construed self is in focus and where perception of tension release through the instrumentality of others is immediate. The neurotic patient who anxiously says: "I can't seem to do the right things to make friends," has as the referent for the I an inference whose structural organization places the empirically-derived introjecting-extrojecting self in the focus.

Or finally, when a college student says: "When I stand up in front of a class, I imagine myself as a leader or a great teacher and I behave accordingly," the referent for the I is the substructure designated as the social self which differentiates roles. The basis for these variants in cognitive organization is to be discovered in the conditions in early childhood which lead to the formation of over-developed boundaries of particular empirical selves.

This set of suggestions is illustrative of the general proposition that the properties of the substructures are determined by the total interbehavior field. And the total interbehavioral field includes not only perceptions referable to objects in the external world but also perceptions referable to the self.

AN HYPOTHESIS ARISING FROM THE THEORY

Here is one hypothesis that can be formulated from the foregoing paragraphs. When studied in terms of reference to the self, the inmates of a school for delinquent boys will fall into two classes: (1) those whose cognitive organizations are dominated by the primitive construed self (S_3), and (2) those whose cognitive organizations are dominated by the introjecting-extrojecting self (S_4). In both groups the social self (S_5), as a cognitive substructure, is absent or poorly developed, probably as the result of overlearning at earlier stages due to personal-social factors. Members of Group 1 respond to imbalance in the interbehavioral field at the level of the primitive construed self: immediate motor discharge (acting out) upon perception of a stimulus object or event. In Group 2 the cognitive structures of the members are dominated by the *Anlage* of the social self—the introjecting-extrojecting self—where there is some time-binding and tension-binding. At this level of self-or-

ganization, discrete acts of others are introjected, such as "Mama says no, no" or extrojected. The delinquent act committed by a member of Group 2 is usually incongruent with the child's inferences about self which are at the level where acts are differentiated in terms of approval-disapproval. A kind of rudimentary conscience is apparent.

The delinquent acts performed by Group 1 are consistent with the level of the inferred self whose main substructure is the primitive construed self. Members of Group 2 are able to evaluate acts in terms of parental approval-disapproval because of the strength of the introjecting-extrojecting self, and cannot easily take the role of the delinquent. From these sentences we can state a high probability that members of Group 1 have a poor prognosis and will turn out to be recidivists, while members of Group 2 have a better prognosis and will become acceptable citizens. That is to say, they may develop a social self—which means skill in taking the role of the other, time-binding, tension-binding, etc. It is predicted that members of a third group, non-delinquents, will be characterized by inferences about self at the level of the social self. These boys do not take the role of the delinquent, nor do they commit serious delinquent acts, as do members of Groups 1 and 2, respectively. How can we test this in terms of the theory of self-organization? We begin with the question: what is the referent for the term "I"? We set up subtle tests designed to get at descriptions of the self in terms of time-binding characteristics, reliance on magic, failure to take the role of the other, physiognomic and literal perception, and other characteristics which are derived from the epistemogenic theory. My colleague, Harrison G. Gough, has advanced a similar theory (13) and has developed a scale of 50 self-reference items which

are answered by the subject. On the basis of these and other descriptions of self, a behavior analyst can infer which substructures of the self are dominant and from these inferences he can formulate predictions of behavior. A study along these lines is now in progress.

Recapitulation. In this essay I take the position that a more careful and sophisticated analysis of the concept of the self (or ego) will help reduce the confusion in personality theory and clinical practice. The description of the self in terms of cognitive structures takes advantage of the most advanced conceptual schemata currently available (18). Because it leans so heavily on cognition, I have labeled the theory "epistemogenesis."

Through interaction with stimuli at different times in the maturational and personal series, various empirical selves or cognitive substructures are organized. These substructures may have different properties of strength and breadth, thus accounting for some aspects of differential conduct. The referent for the "I" or Pure Ego is the cross-section in the present of the total cognitive organization that embraces those more-or-less enduring substructures, the empirical selves.

That the theory is not footless is demonstrated by informally deriving an illustrative hypothesis. The predictions from the hypothesis are testable by means of a set of experiments designed to estimate the "I's" or current self-organization of at least three classes of persons. The experiments are in progress and will be reported along with the fuller description of the epistemogenic theory of the self.

REFERENCES

1. ALLPORT, G. *Personality, a psychological interpretation*. New York: Holt, 1937.
2. —. The ego in contemporary psychology. *PSYCHOL. REV.*, 1943, 50, 451-479.
3. BALDWIN, J. M. *Social and ethical interpretations in mental development* (3rd ed.). New York: Macmillan, 1902.
4. BAYLEY, NANCY. Mental growth during the first three years. *Genet. Psychol. Monogr.*, 1933, 14, No. 1.
5. BUHLER, CHARLOTTE. *The first year of life*. New York: Day, 1930.
6. CALKINS, MARY W. The self in scientific psychology. *Amer. J. Psychol.*, 1915, 26, 495-524.
7. CHEIN, I. The awareness of the self and the structure of the ego. *PSYCHOL. REV.*, 1944, 51, 304-314.
8. CLAPARÈDE, E. Note sur la localization du moi. *Arch. de Psychol.*, 1924, 19, 172-182.
9. COOLEY, C. H. *Human nature and the social order*. New York: Scribner, 1902.
10. FRENKEL-BRUNSWIK, ELSE. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personality*, 1949, 18, 108-143.
11. FREUD, S. *A general introduction to psychoanalysis*. New York: Boni and Liveright, 1920.
12. GESELL, A. *The mental growth of the pre-school child: a psychological outline of normal development from birth to the sixth year, including a system of developmental diagnosis*. New York: Macmillan, 1925.
13. GOUGH, H. G. A sociological theory of psychopathy. *Amer. J. Sociol.*, 1948, 53, 359-366.
14. HALL, G. S. Some aspects of the early sense of self. *Amer. J. Psychol.*, 1898, 9, 351-395.
15. JAMES, W. *Psychology, briefer course*. New York: Holt, 1915.
16. JERSILD, A. T. *Child psychology*. New York: Prentice-Hall, 1940.
17. JUNG, C. G. *Two essays on analytical psychology*. New York: Dodd, Mead, 1928.
18. KRECH, D., & CRUTCHFIELD, R. *Theory and problems of social psychology*. New York: McGraw-Hill, 1948.
19. LATIF, I. The physiological basis of linguistic development and of the ontogeny of meaning. Part I and II, *PSYCHOL. REV.*, 1934, 41, 55-85, 153-176.
20. MCCARTHY, DOROTHEA. Language development in children. In *Manual of child psychology* (L. Carmichael, Ed.). New York: Wiley, 1946.
21. McDOUGALL, W. *An introduction to social psychology*. Boston: Luce, 1921.

22. MEAD, G. H. *Mind, self, and society*. Chicago: Univ. of Chicago Press, 1934.
23. MUNN, N. Learning in children. In *Manual of child psychology* (L. Carmichael, Ed.). New York: Wiley, 1946.
24. PILLSBURY, W. B. The ego and empirical psychology. *Phil. Rev.*, 1907, 16, 387-407.
25. PRATT, K. C. The neonate. In *Manual of child psychology* (L. Carmichael, Ed.). New York: Wiley, 1946.
26. SARBIN, T. R. Contributions to role-taking theory. I: Hypnotic behavior. *PSYCHOL. REV.*, 1950, 57, 255-271.
27. —. The concept of role-taking. *Sociometry*, 1943, 6, 273-285.
28. —, & FARBEROW, N. L. Contributions to role-taking theory. II: A clinical study of self and role. (In press.)
29. SHIRLEY, M. M. *The first two years: a study of 25 babies, Vol. II, Intellectual development*. Minneapolis: Univ. of Minnesota Press, 1933.
30. SKINNER, B. F. "Superstition" in the pigeon. *J. exp. Psychol.*, 1948, 38, 168-172.
31. STERN, W. *Psychology of early childhood*. New York: Holt, 1924.
32. SYMONDS, P. M. *Dynamic psychology*. New York: Appleton-Century-Crofts, 1949.
33. TITCHENER, E. B. *An outline of psychology*. Macmillan, 1896.
34. WEISS, E. Projection, extrojection, and objectification. *Psychoanal. Quart.*, 1947, 16, 357-377.
35. WERNER, H. *Comparative psychology of mental development* (rev. ed.). Chicago: Follett, 1948.

[MS received November 17, 1950]

STRUCTURAL RIGIDITY IN RELATION TO LEARNING THEORY AND CLINICAL PSYCHOLOGY

BY RAYMOND B. CATTELL

University of Illinois

AND

ALVIN E. WINDER

Roosevelt College

I. OBJECTIVES OF THE DISCUSSION

The time has become ripe for integrating the concepts of rigidity which have grown up, with undue isolation, in the all too widely separated experimental and clinical fields. By the standards of pure science it cannot be denied that on the one hand *clinical* "hypotheses" using the word "rigidity" have been slovenly in conception, unstable in application and retarded in the development of appropriate crucial tests. The experimental conceptions have been precise and have yielded a harvest of extensive investigation (1, 2, 3, 4, 31, 33) which has recently culminated in some definite structuring of the field and the clarification of past obscurities (7). Nevertheless, it is possible that some rigidity phenomena in the clinical field have not yet been brought within the realm of experiment. Certainly some wider experimental exploration remains to be done before the whole area of interest can go forward in more intensive, interrelated and well-conceived research.

The object of this paper is first, to survey the clinical literature in order to inventory and classify the specimens of alleged "perseveration" or "rigidity," and, second, to bring the various conceptions of rigidity into an orderly theory from which certain crucial experiments may be suggested. The exposition will not follow quite that order. Instead we shall first build up a clear

theoretical position on the basis of available general and experimental evidence, then consider clinical concepts in the light of this structuration, and finally proceed to ask what further experiments are required to clear up any inconsistencies.

II. A THEORETICAL ANALYSIS OF THE FACTORS IN OBSERVED "RIGIDITY"

A survey of the most common uses of rigidity shows that it refers to a *failure to "adapt," by the use of a shorter behavioral route than the usual one to a given goal, when circumstances make the shorter route possible*. The reader will note that this excludes the phenomenon of "process rigidity"¹ which for a short phase in the history of perseveration investigation was included in rigidity. It will be recognized that rigidity as defined above is nothing more than *slowness of learning* when the conditions for learning exist. Let us first see what different classes of phenomena fall *within* this definition before we proceed to ask whether there are any usages of rigidity which remain outside the definition.

It must be recognized at the outset that there are two general situations in

¹ By process rigidity we mean the mere temporal continuation, as by a momentum, of the immediate conscious or behavioral process itself. This meaning may be set aside from the *structural* concept of rigidity, i.e., rigidity in connection with change in capacity to react.

which the word rigidity has been applied to failure to adapt or learn: (1) failure to achieve (even once) the shortest path in a new situation, and (2) failure to *acquire a habit* of taking the shortest path in repeated presentations of the situation. Obviously (2) depends partly upon (1) and brings in addition some effects in the realm of retention and extinction.

A previous theoretical and experimental discussion (7) has presented evidence for an analysis of factors in rigidity in the above two situations of the following nature. It should be recognized that this cataloguing always assumes forms of rigidity to be distinct, unless correlation has proven them to be the same, i.e., initially it *may* start with more forms than actually exist, and then proceed by experiment to verify or reduce the hypothetical variety.

A. In New Situations

1. Defective perception of relations. Here the problem solving fails through insufficient "g" or deficiency of more restricted perceptual abilities.

2. Defective production of possible responses upon which either relation education (insight) or blind reward might operate. This covers (a) *reproduction* of responses made elsewhere before and perhaps also (b) *creation* of new responses.

a. Even on *a priori* grounds it can be seen that reproduction is affected by a variety of factors, although our analysis of the actual number must be contingent upon further experiment. However, the following main categories can safely be considered. i. Defective breadth of previous experience—an environmental, historical influence. ii. Defective powers of retaining experience. By this is meant power of memory as distinct from deficiency in the powers of perception or the dynamic influences

involved in committing to memory. As demonstrated through the volume of recall when operating on some loosely defined mental set this may be identifiable with the well-known "fluency of association" factor (2, pp. 420-428). iii. Possibly defective orderliness in committing to memory and in consolidation, which may be Thurstone's M factor (32). Abilities as in A.1. above also obviously determine committing to memory.

b. Defective spontaneity of new responses. Spearman's *Principles of Cognition* (31) considered the appearance of cognitive content to be due to *anoegetic* (reproducing) and *noegenetic* (novelty creating) laws. Relation or correlate education, however, was considered the sole source of noegenesis. But it is possible that the appearance of entirely new responses, in the "random" behavior of the strongly deprived animal in the puzzle box, may be due also to a further process which must be considered *sui generis*. New responses in this further sense presumably are mixed with a wider array of reproduced responses, in the *dispersion with continued deprivation and excitement* in Cattell's third dynamic principle (9, p. 635). Probably one would have to consider three factors which also operate in other processes, namely, length of deprivation, strength of drive, and excitability of the individual. (It should be understood that length of deprivation is considered quite distinct from strength of drive.)

3. Defective strength of motivation may operate as just stated but presumably operates also in regard to education of relations (1) above, though intelligence test scores respond relatively little to motivation changes, and to recall, in the process of reproduction. The factor of fatigue must be considered here as operating to reduce the effective strength

of motivation. We must also distinguish between the effect of strength of motivation at the time of recall and

4. The historical fact of defective motivation at the time of initial perception and committing to memory; the reaction which the individual makes to some recognized part of the situation being too rigid to be supplanted by a possible more adaptive response. This is a rigidity of established as distinct from new connections.

5. Immediate disposition rigidity, i.e., true rigidity classically considered an inability to modify a response despite application of standard degrees of reward and punishment.

We have next to consider that kind of rigidity, i.e., failure, which is illustrated in *repeated* adaptive response situations.

B. In Repeated Situations

Here we are dealing with the typical learning curve in a situation exactly repeated. The adaptation which occurs during the actual time of exposure to the situation is adequately accounted for by the A analysis, above, and the question here is whether there is also some *differential readiness to reinstate at the second exposure* successful responses already made at the first.

Conceivably the process of retention between exposures is determined by factors different from or operating differently from those in the immediate learning situation. It will be understood that we refer here to the retention of the particular adaptive responses to the particular situation, not to the retention of a general pool of available responses as in A2(a)ii above.

Learning theory has given much attention to this question. Hull's (18) systematic investigation of this problem treats the effects of repetition in

terms of the postulate of "reinforcement" learning and states that the maximization of habit strength is dependent primarily upon the number of reinforcements. The relationship between effective habit strength and drive in causing behavior has been suggested by Perin (27, 28) by using groups of rats in the Skinner box learning situation. Perin determined the resistance of a learned habit to extinction as a joint function of (1) the number of reinforcements of habit strength and (2) the number of hours of food privation (drive) preceding extinction.

Experimental extinction is but a special case of the tendency of all responses, through the fact of their occurrence, to develop a barrier to their own repetition. This inhibition wears off in time and leads to the spontaneous recovery of response noted in conditioned reflex terminology. The available evidence suggests that:

1. Interest (motivation) during the interim (consolidation) period is of importance. We should also consider as a factor the "fading out" of habits in addition to weakness of interest, namely some tendency to revert to the status quo, as postulated in 2 below.

2. Latent disposition rigidity, distinct from the immediate rigidity of A5 above, which operates to restore older responses vis-à-vis newly learned responses during the latent periods between learning exposures. This latent disposition rigidity is presumably a constitutional tendency adjusted by genetical selection to the expectations set by the physical environment. That is to say there is a certain probability of the environment remaining unchanged from occasion to occasion and an optimum retention rate of learned habits in relation to that probability.

III. AGREEMENT OF ANALYSIS WITH EXPERIMENTAL EVIDENCE

Analysis of the factors in observed, apparent, operationally defined "rigidity" or, more precisely, *poorness of learning*, shows that it is pointless to use the precise term rigidity for all the phenomena of defective learning, much of the behavior in which is accounted for by other known influences, and we have accordingly limited the final use of the term rigidity at the present to the postulated processes A5 and B2.

In summary it will be seen that the analysis describes exhaustively a number of independent processes in learning and takes the most likely hypothesis about the factors determining each. Some of these factors operate at once in two or more different processes but most are peculiar to one process. Further, the analysis indicates that if "poorness of learning" (or of problem solving or adaptation) is measured in a single, immediate situation, we should expect the collection of factors to be different from, though overlapping with, those for "rigidity" when measured by degree of adaptation to a repeated situation.

These collections of factors cover influences (1) in the situation, (2) in the temporary state of the subject, and (3) in the more permanent individual makeup of the subject. It would be possible to simplify the problem of holding one or more classes of factor constant and omitting them from discussion. Holding factors constant may be a satisfactory technique when applied to the problem of experimental design; however, it is initially desirable to survey them all. They may be summarized as follows:

A. New Situations

1. Ability, operating in the actual situation (A1) and in the provision of memories (A2(a)iii).

2. Experience, in A2(a)i.
3. Power of retention, in A2(a)ii.
4. Skills in committing to memory, in A2(a)iii.
5. Provocation to excitement of the individual, in A2(b).
6. Excitability of the individual, in A2(b).
7. Strength of immediate motivation or freedom from fatigue, operating in A1, A2(a), and A2(b).
8. Strength of past motivation, operating in A1, A2(a)ii, and A2(a)iii.
9. Immediate disposition rigidity, operating in A1, A2(a)ii.

B. Repeated Situations

- 1, 3, 4, and 7 from A, plus
10. Interest, operating in the interim, B1.
11. Latent disposition rigidity, operating in B2.

It is possible, of course, that these logically separable eleven factors are in some cases physiologically or psychologically connected, notably in that immediate and latent rigidity may turn out to be the same; excitement with deprivation and strength of immediate motivation might conceivably be one; and excitability might prove to have some relation to rigidity. Our design, as indicated above, has been to assume that they are distinct until proved the same.

Let us see how far this analysis is supported or modified by available experimental evidence.

The role of 1, ability, in what has sometimes been called "rigidity" but which we here call *poorness of learning* or *adaptability*, has been overwhelmingly recognized and experimented upon. It has been recognized in the "rigidity" of mental defectives and brain-injured individuals. Neurologists have long recognized and referred to the clinical picture of rigidity presented by the brain-

injured individual. Goldstein particularly has sought to reduce these clinical data to a consistent definition of rigidity. Piotrowski, Oberholzer, Weigl, Harrower and others have shown through the use of the Rorschach and other tests that perseveration and lack of plasticity are characteristic of the brain-injured. Goldstein (15) has included under his concept of rigidity the feeble-minded, intoxicated and schizophrenic, as well as the brain-injured. He has hypothesized two types of rigidity, primary and secondary. Primary rigidity is due to impairment of higher mental processes, a basic lack of ability to change set. Secondary rigidity is due to impairment of abstract thinking. When the task is too difficult the subject avoids the catastrophic response by continuing to do a task that he has solved before.

Lewin (23), Kounin (20, 21) and Werner (34) have all insisted on rigidity as an important characteristic of feeble-mindedness, but they failed to show by factorization that the phenomena they described were not due to low "g" itself. Their writing consequently employs rigidity in the broader rather than in the more restricted of the two senses we used above. Cattell (3, 4) has pointed out that the long line of researches on "classical perseveration" recognized at an early phase the existence of a negative correlation between their tests and intelligence and so designed them as to eliminate intelligence, demonstrating that intelligence and the factor in their tests were independent factors in general "rigid" behavior. Recent work (16, 24, 25), which has examined rigidity from the problem-solving approach, has done the same. Finally, the recent integrating study of rigidity in many types of situations has demonstrated by factor analysis that low intelligence is one major

factor (8) in a wide range of such behavior.

The roles of past experience and power of retention have never been investigated, but some researchers have endeavored to hold them constant. Fluency of association has recently been demonstrated to be a powerful (negative) factor in many manifestations of rigidity (7) and it is conceivable that high fluency score is a direct expression of greater available memories, owing to better retention. The role of skills in committing to memory (No. 4) is well known through research in educational psychology. But Nos. 2, 3, and 4 cannot at present be brought into very positive relation with measurements of rigidity to produce any simple laws.

There is no adequate experimental work relating rigidity vs. flexibility to length of deprivation (dispersion through excitement) No. 5, or to excitability of the individual No. 6. However, as we shall see when studying the clinical contribution below, there is some evidence of flexibility with higher general emotionality.

Concluding the above list of factors, we encounter influences No. 7, immediate motivation; No. 8, past motivation; and No. 10, interest strength during the interval between learning, all of which have received considerable attention in experiments on learning. Their roles in learning theory are too well known to require extensive discussion, but, seen from the viewpoint of "rigidity as failure to learn," they can be summarized as follows.

Immediate motivation operating at the time of the initial perception may cause the reaction the individual makes to some aspect of the situation to be too rigid to permit the creation of an adaptive response. In the case of past motivation operating through the process of recall the individual may react to some recognized part of the present

situation very similarly to the way in which he reacted to it in the past. This would have the effect of blocking the creation of a more adaptive response. Interest strength during the interim may act to maintain rigidly a previous learned response so that it is available as a fixed solution to any situation that it can be made to fit, and in this way block the creation of a more adaptive response.

This conclusion of the survey leaves us with one true rigidity, which we shall call disposition rigidity No. 9, and also a possibly distinct latent rigidity, No. 11. The latter we shall set aside for the moment, since it requires oblique and questionable inferences from inadequate researches. The remaining "rigidities" are really motivation or ability factors.

Although it is true that research has been very active concerning this "immediate (disposition) rigidity" in the first learning situation, it nevertheless wasted some of the early inquiries in pursuing the false scent of what has since been set aside as "process rigidity." The history of development since Neisser, Muller, Heymans, and Wiersma, Gross *et al.* has been summarized elsewhere (3), and it will be seen that Spearman, following these initiators, considered perseveration, p-factor, or rigidity as a kind of dynamic momentum. Thus it would be expected to reveal itself by the relative slowness of an operation requiring rapid oscillation between processes A and B, when compared with steady repetition of A or B separately.

Fortunately the line of researches begun by Spearman were equipped with a method—factor analysis—which was bound sooner or later to reveal the emptiness of a false conceptualization and, in 1935, Cattell (1) pointed out that two distinct principles had been involved in perseveration tests and that

a clear-cut general factor existed only in those tests which had not been included in the main battery! Thus it was found that the essence of perseveration or rigidity existed only in creative effort tests, measuring difficulty in departure from an old habit, and that "alternation" tests, insofar as they have anything in common, apparently have only an accidental contamination with the disposition rigidity factor. This was brought out soon after by the work of Walker, Staines, and Kenna (33). Incidentally it has been shown again in the very recent and neat investigations of Kleemeier and Dudek (19) on rigidity as flexibility (undertaken without apparent awareness that this conception had been tried and rejected earlier) that factorization yields only the well-known abilities and no factor of process rigidity as such.

The recent experimental study by Cattell, Tiner, and Winder (7) set aside process rigidity from the beginning, but otherwise took as wide a sampling of alleged rigidity situations as possible in seventeen test situations. Markers were inserted for intelligence and fluency, in accordance with the theoretical analysis, which was a rough beginning of that made above, and an especial attempt was made to see whether the creative effort (3) type of test identified in motor manifestations would also manifest itself in perceptual performances. An attempt was also made to include some clinical-type perseveration.

The results showed that the theoretical analysis was largely correct and that a good deal of the variance in rigidity test situations is accounted for, first, by the independent and long known factors of intelligence and fluency. The remainder was due largely to two factors of rigidity, the first containing the markers for the classical rigidity factor of many past researches and the second manifesting itself in

what has been contingently named ideational inertia (poor ability on riddles and reconstructing hidden words). The interesting finding in regard to the classical factor is that it *does* include rigidity of perception (and slow speed of flicker fusion) as well as the longer attested motor rigidity.

Now, as indicated, the nucleus of tests in the classical factor has been *creative effort motor* tests, e.g., drawing the letter *e* repeatedly, beginning at the opposite end from the usual, *in which the subject sees clearly what he has to do and wills to do it, but cannot suppress older motor responses to the situation and install new ones.*² The learning is of the A or immediate type above (except that some consolidation may occur during the few minutes of the test), and no experiment has been carried out to see whether the factor is the same as that which might be present in the B type of situation—latent rigidity over a long interval of retention.

A basic theory has been put forward to account for the indubitable and provocative connections found between this classical "p-factor" and measurements of personality as summarized elsewhere (4). This theory of personality connections is relevant here in that it makes rigidity a cause of defective volition. On experimental grounds one might entertain the possibility that individual differences in performance on tests of the above type could arise from differences in strength of motivation (No. 7 above) and not require any supposition of differences in rigidity. At least it might be due to differences in strength of will (the organized ego) as contrasted with independent habit

systems. The experimenter may reject this explanation on the grounds that he has attempted to induce maximum and equal motivation in all subjects. But even if he is not successful in this argument the theory mentioned here would reconcile the two viewpoints by saying that, with the same incentive, the individual who can bring the stronger willed (ego) action is he who has lower original rigidity. The association of weak ego strength with high rigidity is a fact, but to say the former derives from the latter is an hypothesis, and must be considered as such in any experimental designs to investigate p-factor more profoundly.

The contribution of the Cattell, Tiner, and Winder study to the understanding of this classical p-factor is possibly disconcerting to the above theory in that one of the two highest tests involved looking for hidden figures in a given picture. However, although one cannot "will" to see the new figure here, since one does not know what it is, one can will *not* to see the habitually seen figure; so the generalization may remain that this factor is one of resistance of existing responses to newer responses which are explicitly known to be the more successful ones. That is to say, it is a "habit vs. will" rigidity, when intelligence, reproductive powers, strength of motivation, fluency of random responses etc. are held constant.

With the restriction of the term rigidity to experimentally defined and investigated entities, namely the classical rigidity factor (motor and perceptual) and the new factor of "ideational rigidity," and with the further, hypothetical division of the field of experimentally observable "non-adaptation" into ten remaining, tentatively-defined factors, we may next examine the use of rigidity in the clinical field.

² It is possible that although the subject sees clearly what he has to do, so that no relation education or reproduction appears to be involved, some "fluency" in providing suitable motor responses for the given, envisaged behavior is also required.

IV. CONCEPTS OF RIGIDITY IN THE CLINICAL FIELD

Our purpose is to determine how far any reasonably apt uses of the concept of rigidity in clinical observations resolve into the above, or, alternatively, require tentative additional concepts.

The outstanding instance of rigidity in the clinical domain is what has been called the "neurotic paradox," in which, meeting repeated punishments, the individual apparently defies the law of effect by repeating the maladaptive behavior. If this were the true explanation, the phenomenon would lie outside the rigidities just discussed, in all of which the law of effect is supposed to be operating and in which the failure arises from other causes.

Setting aside the substantial fraction of cases in which real but masochistic satisfactions of the law of effect explain the behavior, we are left with two known mechanisms: (1) the repetition compulsion, which Freud interprets as a satisfaction of a mastery, aggression or death drive, and (2) the presence of dissociations caused mainly by repression, which results in the individual being unable to connect the punishment with the behavior which causes it. Clinicians may wonder at "mainly" in the above sentence, but one must admit other possible factors in the breakdown of the law of effect. For example, a rat taught to press a lever and wait three seconds before eating the food which drops will generally fail if the interval is stretched to six seconds. He will then either take the food prematurely, and be repeatedly punished, or starve. Here, as in many clinical "perseveration" examples, the rigidity is nothing more or less than the undying persistence of appetite which has lost contact with adjustive mechanisms—it is the demanding *id* without the directing ego. Consequently the whole class of influences which limit

control of impulse are factors in rigidity. Many of these, e.g., intelligence and retention, have been discussed above, but we may need to add here "temporal span of integration," meaning thereby the above limiting factor in the individual's ability to retain a mental set regarding conflict between drives, as well as such factors as fatigue, alcohol, and various physiological determiners of efficiency of control.

It is possible that some limiting factor of this kind is involved in the next broad clinical instance of rigidity to be considered—that in schizoid or schizophrenic individuals. Even in the normal range of temperament it is striking that behavior ratings of rigidity have their highest loading (in three studies, see below) in factors A and H—schizothymia vs. cyclothymia. Rigidity is defined in these ratings as "always doing things in one particular way; non-plussed when his routine is upset; following the letter rather than the spirit of instructions." Since the factorization shows that intelligence is not involved, we perhaps have to deal with that inability to maintain an over-all mental set which Rodnick and Shakow (29) have found so conspicuous in schizophrenics. Our hypothesis will be that although other factors are present, there is a special limitation causing this kind of rigidity, namely, an inability to maintain massive mental sets, which may be identical, in nature if not in causes, with those in the rat example above, as well as in fatigue and in physiological impairment. Thus, although the H-individual is not devoid of the factors necessary to perceive the approach necessary to solution, he is incapable of maintaining the approach; for the general set disintegrates into rigid, trivial component sets which cannot be brought adaptively into relation one with another.

A form of rigidity frequently consid-

ered in clinical discussions is essentially the same as that studied experimentally as an *Einstellung* (mental set) effect by Guetzkow (16), Luchins (24, 25), Ro-keach (30), Christie (11), and others. Here the individual is first taught to solve a problem in a certain way, and then is given a roughly similar problem for which the old habit is inapt. Although this is, in a broad sense "operational," i.e., a single operationally-defined phenomenon, such failure to solve the problem could occur from several of the basic factors we have considered above, e.g., low intelligence, low fluency, the classical perseveration-rigidity factor, and so on. Actually the work of Christie (11) gives correlations which suggest that the classical rigidity factor is substantially involved in the *Einstellung* situation while the work of Laycock shows intelligence to have a substantial role in intelligent transfer.

A fruitful initial approach to studying rigidity in relation to the total personality or clinical picture is to follow the findings of factor analysis on the contribution of various personality factors to *rated* rigid behavior. As indicated, source trait A (and its relatives, H and F) account for the consistently largest slice of the variance, but three studies (2, 5, 6) agree also in assigning some fraction to source traits D and E, while some is also involved in M and G. That of E (dominance, egotism) can be dismissed as a social rigidity which is rigid or maladaptive only by the standards of a different social goal.

Similarly it is clear that the rigidity which some clinicians see in G (positive character integration) is largely a value judgment. Here the individual rigidly and obsessively maintains certain ethical standards and practices which are as great a deviation in one direction from the average adjusted person as the delinquent is in the opposite direction. It is perhaps only a passing psychiatric

fashion to regard this deviation from the statistical norm as abnormally rigid (confusing the two senses of the word "normal") (2, p. 15), but it should be mentioned as part of current discussion. Essentially, however, the individual in this case refuses to adapt to an immediate temptation or opportunity because it would interfere with more remote and obscure satisfactions. Though it may appear when viewing a narrow sector of behavior operationally as rigidity, it is in no psychological sense rigidity when the total reward situation is concerned; indeed it is the converse of the rigidity of poor control.

The consideration of factor M raises the difficulty that its nature is not yet fully understood, but it may be approximately defined from the rating picture as a worrying, anxious, over-inhibited conventionality—perhaps that which some psychiatrists have been confusing with the "character" factor G above. There is considerable evidence, both in animal experimentation and clinical cases, that fear or anxiety of certain kinds increases rigidity of behavior. G. Hamilton, Maier, Krechevsky and J. Hamilton have reported animal experiments in which persistent non-adjustive behavior is the result of an emotionally charged situation. J. Hamilton and I. Krechevsky (17) have demonstrated that if rats are trained to find food at one arm of the T-shaped alleyway and then given a severe electrical shock while running to the food, it becomes very difficult to eliminate the habit of turning in the first learned direction when food has later been shifted to the other arm of the T.

The clinical evidence with regard to this point has been reviewed by Noyes and Fenichel. Noyes (26) finds that phobic reactions, tics and compulsive movements are stereotyped anxiety avoidance reactions common to many behavior disorders. In summing up the

clinical observations of the psychoanalytical school on this point Fenichel (13) says that in the face of neurosis the ego acts to prevent further breakdown by adopting secondary adjustments that involve a curtailment of flexibility and freedom. Adjustments of this kind are encountered in behavior patterns through which dangerous situations are avoided or reassuring ones induced. In extreme cases the rigidity is a total one; in less extreme cases a relative elasticity may be preserved so that the rigid pattern becomes pronounced whenever anxiety is felt, and is somewhat relaxed when an experience of reassurance permits the individual to ease the barriers. All ego-defense mechanisms produce some degree of unadaptive rigidity.

Fisher (14) in a recent study of rigidity in a normal and abnormal population, based on a series of personality tests, has summarized his findings in terms of the adjustment viewpoint. He suggests that the amount of rigidity present in a subject should be viewed in terms of a ratio of rigidity shown when dealing with non-threatening tasks as opposed to rigidity shown for threatening (ego-involved) tasks.

How are we to conceive this rigidity? Our hypothesis will be that fear and anxiety, perhaps in special conditions only, act specifically to restrict the range of random behavior (the latter emanating either from reproduced past possible reactions or from new "dispersion" behavior). Incidentally this should be distinguished from the rigidity of the obsessional neurotic, which happens to arise from undischarged anxiety or guilt, but in which the essential cause of the rigidity is the dissociation, which would produce lack of adaptation in any impulse being discharged. In sum, both anxiety and rigidity, as clinically rated, fall largely in personality factors A —, F — and M — (except for

the general emotionality of C —, below). Of these the first is most important, seemingly corresponding to a measure of the individual's total frustration (and therefore hostility and anxiety). The well known rigidity of the rat in Maier's experiment with continued choice failures is available for investigation as a classic form of this A — rigidity.

The discussion of the last of the personality factors known to have consistent correlation with rigidity, namely D, involves us in the general question of the role of emotionality. D, contingently labelled infantile emotionality vs. frustration tolerance, has highest loadings in "demanding," "impatient," "attention-getting," "self-willed," and "uncontrolled." Both in Burt's studies of emotional responses in children and Cattell's studies of adults it clearly factors out as distinct from the dimension of general emotionality (general neuroticism). Hunches as to the nature of this factor have ranged from mania and hysteria syndromes to simple lack of maturation, i.e., emotional infantilism. Whatever its nature may be, it is definitely associated with rigidity of impulsive responses, which are not varied or adapted. Now we must recognize here that although most instances of rigidity are failures to take the best path to a desired goal, the social and clinical evaluations bring out instances not found in the experimental survey in which no possible satisfaction of the goal exists and where adaptation therefore consists in complete suppression or repression of the impulse. It is possible, and we shall adopt it as a firm hypothesis for experimentation, that the rigidity of factor D is that kind which arises from inability to suppress or repress drives which cannot obtain satisfaction.

Although general emotionality (the negative direction of C factor) is non-

linearly, positively correlated with the classical rigidity factor (p-factor) as numerous researches summarized by (4) show, there are nevertheless some aspects of rigidity in different and wider senses which seem negatively related to general emotionality. This has been brought out most clearly in a recent clinical P-technique study (8) where it was concluded that "some kind of energy, fluency, and novelty of attack are by-products of the instability and impulsiveness of low C" (i.e., high emotionality). Specifically the tests which increased with low C rating were fluency of association, daily fluctuation of attitudes and ability to carry out multiplication with various changes of symbol meaning.³ It is reasonable, indeed, to expect that lower (low C) stability would result in more varied activation of memory material than would arise in a highly integrated personality where the paths of action are clear cut. Consequently fluency and variety of response (mostly incorrect for the given situation in its normal form, but containing some novel, suitable variant response when the situation is suddenly modified) would be higher with low C.⁴

This same instability and lack of internal integration would also reduce rigidity in a way different from any yet

discussed. For now the very *goals* are likely to alter despite the stimulus situation remaining constant. The perception of and accompanying attitudes toward a given situation alter because the individual's appetites alter, because other satisfactions are blocked, and because the loose integration of the individual's dynamic lattice (2, 9) has resulted in certain satisfactions being obtained through new paths. Presumably the obtaining of new satisfactions through old situations and responses necessarily brings about the obtaining of old satisfactions through new situations and responses—or at least the attempt to do so. Such instability is therefore likely to reduce rigidity in the sense previously defined, i.e., as the inability to discover a shorter path to satisfaction; but it also brings to our notice two new definitions of rigidity outside the definition as rigidity-in-approaching-a-given-goal.

First we have here a reduction in rigidity in reaching *any* given goal due to a looseness of the whole dynamic structure. The events in any given goal path are recognized as being dependent on repercussions from the whole of the dynamic lattice. Secondly, we have a possible rigidity or flexibility of the final goals themselves—an "ergic rigidity" which may be distinct from the classical rigidity factor. This last lies outside the whole class of definitions yet considered. Regarding the first we may note that the experimental results clearly demonstrate (fluctuation of attitudes) that day-to-day attitude fluctuation is associated with A + factor as well as C - factor, i.e., the cyclothyme is also given to large changes from mood (appetite) or low integration. This is apparently quite distinct from the form of low rigidity (breadth of mental set) noted above to be associated with cyclothymia vs. schizothymia.

Looking back at the clinical survey

³ The last association is not too well established. One must note also that Eysenck (12) found a reversal of sign of this fluency correlation when one moves into the more extreme range where hospitalized neurotics are involved.

⁴ This may seem to conflict with our earlier argument that the psychoanalytic view of character as something inimical to adaptability is wrong. Actually, however, one must bear in mind that this same instability which provides the apt response in an unusual situation is constantly interrupting apt responses in the usual situation, or in the repetition of the initially unusual situation. It operates against rigidity in the A (initial) situation but in favor of it (ineffective retention of learning) in the B (repetition) situation.

we see that it has indeed introduced new concepts, first in the form of an entirely fresh, additional variety of rigidity—namely, rigidity of goals—and secondly in the form of new factors within goal-path rigidity, mostly factors that have to do with structure rather than powers. Before systematically examining these factors, let us give contingent definition to the new form of rigidity as *goal-rigidity* in contrast to *goal-path rigidity*. This is presumably the rigidity which is specifically overcome in sublimation.

V. EXPERIMENTAL DESIGN

Research at this point needs first to ask the question: "Is the above analysis a correct account of the number and nature of the forms of rigidity?" The analysis rests in some places on *a priori* reasoning alone; in others, on experimental evidence of a general kind and, in regard to three or four factors only, on incontrovertible, specific experimental evidence. Research needs to tighten on the first two classes. But when our understanding of the factors at work is thus adequately structured, there will still remain the question, "What exactly are the properties, incidence, and genesis of each of the factors?" However, even before research is carried much further, we know enough to decide which factors are better designated by some term involving rigidity and which by other terms, and such a sorting is best undertaken before further experiment. In this it is economical to indicate the hypotheses involved by the terminology itself.

The last three sections have been essentially discussions from three different starting points, each leading to an analysis. The analysis has therefore as yet reached only the level of a collection, in which factors of a situational nature, of a constitutional nature, of a goal rigidity, and of a path rigidity type etc.

have been left in unorganized juxtaposition. Let us see how this collection can be organized. First, as to operational concepts of rigidity we can distinguish the following:

A. Process rigidity (effects in rapid alternation of activities); B. Goal-path learning rigidity, (a) in new, immediate situations, (b) in the same situation repeated; C. Goal rigidity (ergic rigidity). This is a resistance of the biological goal itself to change, as opposed to easy sublimation. Whether one should distinguish from this, as a fourth variety, the sheer tendency of an unsatisfied drive to go on demanding satisfaction, i.e., to be indefatigable, is at this juncture a too fine academic point. Some of the rigidity in dominance (E factor) and infantile emotionality (D factor) may be of this type.

Of these conceptions we have set A aside as experimentally undemonstrated, and, moreover, beyond our present interests, while C is perhaps a largely single biological factor in itself. The more numerous factors we have considered are concerned essentially with B, though some may conceivably operate also in C. For ease of relation to experiment the full inventory of factors set out above will be set out separately for conditions B(a) and B(b) above. This final list is developed by the expansion and by the addition of clinically considered rigidities. It is also improved by more systematic definition than was there possible.

Factors in goal-path rigidity.

B(a) In new, immediate situations.

1. Defective perception of relations (abilities).
2. Defective amount of general previous experience.
3. Defective observation and retention of previous experience.

4. Defective ability to maintain a broad, integrated mental set.
5. Defective motivation strength in immediate situation.
6. Lack of "dispersion" of response and of recall under inadequate deprivation.
- 7a. Lack of dispersion through low excitability.
- 7b. Lack of dispersion through low excitement (state).
8. Lack of fluctuation of lattice through high integration C factor.
9. High inhibition due to fear and anxiety (Factor A —, possibly F — and M —).
10. Presence of dissociations, i.e., "structural" causes of rigidity, in specific repressions or through alcohol, fatigue, etc.
11. High strength of existing responses, through much rewarded previous experience, or through inherent strength due to D factor.
12. High immediate disposition rigidity, inherent in the individual, shown in persistence of both responses and perceptions.

B(b) In repetition of same situation.

Here it is assumed that the correct response has once been made and rewarded in the previous situation and that the additional factors determining its nonreproduction, i.e., favoring rigidity, are:

13. Defective strength of motivation in previous situation.
14. Defective skills in committing to memory.
15. Defective motivation in retention interval, including retroactive inhibitions from drainage to other motivations.
16. Defective motivation in recall situation.
17. Defective powers of retention specifically for successful responses,

which might be different from 3 above.

18. High *latent* disposition rigidity, perhaps identical with the rigidity of A — factor, showing itself as resistance to change in intervals between learning sessions.

Naturally any measure of effects in (b) will include all the effects of (a)—unless immediate learning is deliberately subtracted, for the (b) processes act upon the result of the (a) processes and are a further function of them. But, in addition, processes 2, 3, 4 (acting subconsciously), 6, 7, 8, 9, 10, and 11 could be operating *directly* in the "consolidation of learning" in the intervals between learning sessions.

The above summary has rejected the repetition compulsion on the grounds that it is a drive to mastery or aggression, operating by the law of effect in the same way as any other drive. The phenomena of repetition exist, but they are not to be considered rigidity in either of the above senses (B and C), the need to repeat being nothing but a guarantee for the learning of control over a particular situation. The summary has also excluded from the concept of rigidity the view of non-adaptability due to character, or the sacrifice of near to more remote goals,—for this view is wrongly designated as non-adaptability. One form of rated rigidity—that of D factor, the opposite to which is frustration tolerance—is also perhaps not adequately accommodated in the above scheme. It is seemingly a form of strength or insistence of existing responses and is lumped with insistence due to strong past rewards, but may be distinct. However, there is no point in multiplying categories until more is known about D factor.

Finally, in any scheme of this kind, one must face the possibility that a factor will act for rigidity in some circum-

stances and against it in others. The case of C factor (integration vs. emotionality, neuroticism) has already been indicated. Defective integration causes more possibilities of novel reactions to the same situation to be tried; but it also causes more losses of correctly adaptive responses that have been successfully established. Now it can be seen that rigidity itself, as latent but not as immediate rigidity, could operate in both directions. In the more rigid organism both the old response and the newly acquired response will by definition be more resistant to change than in a less rigid organism. It is therefore assumed that latent rigidity is a function of age of response, favoring the re-establishment of older responses in contrast to new and partially acquired responses. One must distinguish also between retentiveness and rigidity, the former being a resistance to supplanting by alternative responses. Naturally this distinction is a possible, hypothetical distinction; experiment may show either that these two processes are one or that they can be distinguished, but the latter assumption is clearer and safer.

Now the above analysis may seem to present a somewhat dismaying plurality of influences. But we may note first, on the theoretical side, that only two of the factors in goal-path rigidity are truly to be called rigidities, the remainder being well-known influences, e.g., intelligence, strength of motivation, power of retention, power of maintaining mental set, already recognized in learning theory, personality study, and clinical psychology. And on the side of experimental approachability, we may note that many factors can be controlled. Indeed in the present approach we can simplify the problem at one stroke, for the sake of demonstration and the planning of an immediate research, by dropping operational cate-

gory C (A is already dropped) and by concentrating on B(a), i.e., on goal-path rigidity in the immediate learning situation, which contains only one putative rigidity, "immediate rigidity." This order of approach is desirable because (a) must be cleared up before (b) can be profitably attacked.

Of the twelve hypothesized factors in B(a), general ability can be set aside as already sufficiently demonstrated in goal-path rigidity (8, 15), sufficiently clarified in its nature and sufficiently controllable to be eliminated in future experiments. Similarly we may wish, at a first concentration on the more tangled issues, to rule out breadth of experience as being obvious and easily controllable. Incidentally, Laycock (22) has demonstrated that mere exposure and attention, before the experiment, to a possible array of adaptive responses (not pointed out as related to the later problem) makes a real but surprisingly small contribution to the reduction of rigidity (22). A third factor which might perhaps be omitted as sufficiently demonstrated, in this case clinically, is the contribution of dynamic dissociations to rigidity—the failure to learn resulting from a failure of connection of reward or punishment with response.

This still leaves a formidable list of nine factors worthy of investigation. Broadly two experimental designs suggest themselves, as follows: (1) To measure a group of persons (or animals) on a single adaptive performance so constructed as to give possible scope for each of these nine influences. Two conditions would be set for each influence—high and low—and every possible combination of the influences would be arranged in a Latin-square type of design. By analyses of variance it could be determined if these hypothesized influences in fact have any influence upon goal-path rigidity, and if so, in which cases. Assuming the minimum, two-

fold, division of each influence this would require 29 categories of experimental condition, i.e., 512, which is inconveniently high. The design might therefore be broken down into two experiments, one with four factors (16 categories) and one with five (32 categories) which would permit perhaps a minimum of ten subjects in each category.

(2) On the other hand, since the analysis of variance design would give us only vague indication of the extent to which the influences, if real, are correlated and overlapping, and no indication at all of how many further factors, if any, exist beyond the influences we have posited, a factor analysis design would be decidedly rewarding. Here we should take, instead of one variable giving scope to all influences, a wide sample of variables, some being deliberately designed to give scope predominantly to one of the supposed influences and some to others. All, however, would have to have in common the stipulated class character of immediate problem solving, to fit the area of research on goal-path rigidity. Obviously, therefore, every test should involve a situation in which the correct response has to be found, i.e., no simple speed test. The resultant factor analyses would indicate the true number of factors at work and, if we are skillful in our choice of variables, will sufficiently indicate the nature of each factor.

From a methodological point of view two or three further points need to be explained. If some nine or ten factors are expected, it would be desirable to have not fewer than thirty variables. Moreover, since the variables are chosen in a specific area (learning) so that there is a possibility of all factors having an appreciable loading in all variables, it is most desirable that a sprinkling of quite extraneous variables be

included to help form clear hyperplanes in the rotation process.

A methodological innovation in this design is the inclusion of variations in stimulus conditions along with variations of performance. This hybrid between controlled experiment and factorization *in situ* has been discussed elsewhere from the standpoint of scientific method (10). The individual differences on certain tests are thus partly due to differences in the subjects and partly to deliberately distributed differences in the conditions. Alongside the score list on any such influenced performance variable there stands a score list on what might be called its "conditional variable," which gives the condition rather than the performance scores. This is to be included in the factorization as an indicator of the experimenter's intended influence, which will help to locate any factor of influence in the performances. Parenthetically, the restrictions of experiment will probably require that conditions be varied only over some half dozen levels, with sections of the population equivalent to cuts on a normal distribution located at each level. The resultant correlations will therefore need a correction for grouping.

Eventually and ideally both of these approaches should be tried. The choice for a given experimenter would depend on his facilities for obtaining the conditions required in each. The present article was originally planned as a theoretical examination of these approaches by Cattell and the experimental pursuit of one of them by Winder, but the latter's work has been temporarily postponed.

VI. CONCLUSIONS

1. The phenomena of experimental, social, or clinical behavior to which the term "rigidity" could be applied with any possible aptness are three: (1)

process rigidity (resistance to alternation)—the opposite to which may be called flexibility, (2) goal-path rigidity—opposite to which is capacity to learn, and (3) goal rigidity (or ergic rigidity)—the opposite to which may be called ergic plasticity or capacity to sublimate.

The last has not been experimentally investigated as to factors involved: the investigation of the first and second yields no evidence of existence of the first as a unitary tendency.

2. If we concentrate on the second—goal-path rigidity—we must distinguish between the situation of (a) immediate learning and (b) repetition of the same learning situation. Some eighteen possible independent factors in rigidity can then be analyzed out; twelve in (a) and fifteen in (b), though inasmuch as (b) is a function of (a) all enter into (b).

Leaving aside the six factors peculiar to the situation (b) and some from (a) as being already well recognized and defined, there remain some nine factors which, by hypothesis, could be experimented upon and detected in the (a) type of experimental situation.

3. Only two of these factors in goal-path rigidity fall outside the list of influences already sufficiently labelled by other terms and located by other concepts in learning theory to be deserving of the specific new label "rigidity." These are *immediate disposition rigidity*—resistance to changing to a new response when the response is provided (clearly perceived) and the motivation adequate—and *latent disposition rigidity*, the tendency for a partially established response, despite adequate motivation for its retention, to fail of retention, relative to the older response to the situation, during the interval between learning sessions.

4. Two designs of experiment—one using analysis of variance and one factor analysis—are suggested for checking the above hypothesis as to the number

and nature of factors in immediate goal-path rigidity.

REFERENCES

1. CATTELL, R. B. On the measurement of perseveration. *Brit. J. educ. Psychol.*, 1935, 5, 76-91.
2. ——. *The description and measurement of personality*. Yonkers: World Book Co., 1946.
3. ——. The riddle of perseveration: I. Creative effort and disposition rigidity. *J. Pers.*, 1946, 14, 229-238.
4. ——. The riddle of perseveration: II. Solution in terms of personality structure. *J. Pers.*, 1946, 14, 329-367.
5. ——. Confirmation and clarification of primary personality factors. *Psychometrika*, 1947, 12, 197-220.
6. ——. The primary personality factors in women compared with those in men. *Brit. J. Psychol. Statist. Sect.*, 1948, 1, 114-130.
7. —, TINER, L. G., & WINDER, A. E. The varieties of structural rigidity. *J. Pers.*, 1949, 17, 321-341.
8. —, & LUBORSKY, L. B. P-technique demonstrated as a new clinical technique for determining personality and symptom structure. *J. gen. Psychol.*, 1950, 42, 3-24.
9. —. *Personality: a systematic theoretical study*. New York: McGraw-Hill, 1950.
10. —. *Factor analysis for the social sciences*. New York: Harper, 1952.
11. CHRISTIE, R. The effects of frustration upon rigidity in problem solution. Paper read at American Psychological Association, Penna. State College, 1949.
12. EYSENCK, H. J. Types of personality: a factorial study of seven hundred neurotics. *J. ment. Sci.*, 1944, 90, 851-861.
13. FENICHEL, O. *The psychoanalytical theory of neurosis*. New York: W. Norton & Co., 1945.
14. FISHER, S. Patterns of personality rigidity and some of their determinants. *Psychol. Monogr.*, 1950, 64, No. 307, 1-48.
15. GOLDSTEIN, K. Concerning rigidity. *Charact. & Pers.*, 1943, 11, 209-226.
16. GUETZKOW, H. Personal communication. An analysis of the operation of set in problem solving behavior. (In press.)
17. HAMILTON, J., & KRECHEVSKY, I. Studies in the effect of shock upon the behavior plasticity in the rat. *J. comp. Psychol.*, 1933, 16, 237-253.

18. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
19. KLEEMEIER, R. W., & DUDEK, F. J. A factorial investigation of flexibility. *Educ. psychol. Measmt.*, 1950, 10, 107-118.
20. KOUNIN, J. Experimental studies of rigidity. I. The measurement of rigidity in normal and feeble-minded persons. *Charact. & Pers.*, 1941, 9, 251-272.
21. —. Experimental studies of rigidity. II. The explanatory power of the concept of rigidity as applied to feeble-mindedness. *Charact. & Pers.*, 1941, 9, 273-282.
22. LAYCOCK, S. R. *Adaptability to new situations*. Baltimore: Warwick & York, 1929.
23. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.
24. LUCHINS, A. S. Proposed methods of studying degrees of rigidity in behavior. *J. Pers.*, 1947, 15, 242-246.
25. —. New experimental attempts at preventing mechanization in problem solving. *J. gen. Psychol.*, 1950, 42, 279-297.
26. NOYES, A. P. *Modern clinical psychiatry*. Philadelphia: W. B. Saunders, 1947.
27. PERIN, C. T. A quantitative investigation of delay-of-reinforcement gradient. *J. exp. Psychol.*, 1943, 32, 37-51.
28. —. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exp. Psychol.*, 1943, 32, 95-109.
29. RODNICK, E. H., & SHAKOW, D. Set in the schizophrenic as measured by a composite reaction time index. *Amer. J. Psychiat.*, 1940, 97, 214-225.
30. ROKEACH, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.*, 1948, 43, 259-278.
31. SPEARMAN, C. *The nature of intelligence and the principles of cognition*. London: Macmillan 1923.
32. THURSTONE, L. L. *Primary mental abilities*. Chicago: Univ. of Chicago Press, 1938.
33. WALKER, E. F., STAINES, R. G., & KENNA, J. C. P-tests and the concept of mental inertia. *Charact. & Pers.*, 1943, 12, 32-45.
34. WERNER, H. Normal and abnormal rigidity. *J. abnorm. soc. Psychol.*, 1946, 41, 15-24.

[MS received November 29, 1950]

THE PROCESS NEED

BY IRVING MALTZMAN

University of California at Los Angeles

I

Within the past few years a large number of experiments have investigated various aspects of need states, relevant and irrelevant, and the drive stimuli presumably related monotonically to them (e.g., 9, 6, 17, 28, 34). In contrast to the experimentation in these areas of investigation, the study of another aspect of motivated behavior has been neglected. This is the study of appetite process needs.

In a paper on "The postulates and methods of behaviorism," Spence (30) distinguished between hypothetical intervening state variables and intervening process variables. A state variable "represents a hypothetical condition or state of the organism which is assumed to have resulted from and is defined in terms of the past interactions of the organism and its environment" (30, p. 76). In contrast, process variables "will represent, not states, but hypothetical, non-observable responses, implicit processes, occurring in the individual" (30, p. 76). It might be added that the process variable may also be distinguished from the relatively permanent state variable by its variation or change as a function of time following certain initial conditions.

Spence makes a further distinction between the process variable and the process law as indicated in his statement as to how reactive inhibition dissipates. "Incidentally, it will be noted that this last relationship (reactive inhibition) is the only one which is similar to the so-called dynamic or process laws of physics. This type of law states or describes the laws governing the changes that occur within a system in time" (30,

p. 79).¹ The writer believes this to be an apt description of the hypothetical intervening process variable. If a process variable is one in which changes occur as a function of time, it is also a process law. The intervening variable is the relationship expressed by the guessed-at function.

The kinds of laws governing need states and process needs may be distinguished on the basis of the extent to which time explicitly enters into the formulation of the law. The kind of lawfulness typically sought in psychology is the historical type of law. At the level of intervening variables this is exemplified by Hull's (12) equation for habit strength. Empirically, it is the familiar learning curve. Plotted along the ordinate are certain regularities of behavior that the experimenter has abstracted out of the total behavior, e.g., the per cent of correct responses. Plotted along the abscissa are successive intervals in time, a series of trials, sequence of stimulus presentations, etc. An empirical equation summarizing the data would not explicitly contain time as an independent variable. Time is an indirect referent insofar as the term "trials" contains as part of its operational definition some reference to a succession of experimental manipulations as antecedent conditions.

A similar type of law would be ob-

¹ Another construct with time as the independent variable is J , the delay of reinforcement (10). It is significant that different authors have interpreted inhibitory potential (24) and delay of reinforcement (25) in motivational terms. According to the view presented in this paper, if these hypotheses are substantiated, the motivation constructs in question would be of the process need type.

tained for need or drive states. The dependent variable is some response criterion. The independent variable may be the ordinal number of hours of deprivation or per cent of body weight loss. Here again the antecedent conditions involve some reference to successive intervals of time. But drive state is not explicitly expressed as a continuous function of the independent variable of time.

The distinction in question may be clarified somewhat by physical illustrations. A familiar historical law in physics is that of magnetic hysteresis. In establishing an alternating current electromagnet a certain amount of energy is repeatedly required to reverse the magnetic flux of the iron. The consequent power loss for each cycle is directly related to the frequency of the cycles, among other things, i.e., to the number of previous "experiences." On the other hand, in a process law, given initial conditions such as the acceleration and initial velocity, distance traveled is a continuous function of time explicitly stated as the independent variable.

In the typical animal experiment secondary as well as primary need or drive states are assumed to be relatively constant during the course of a single trial, at least until reward is obtained. The same may be said for a given state of habit strength. It is believed to be constant between trials or during a trial up to the time of receipt of reward. A constancy of this sort is not exhibited by a process need. It will vary in strength during the course of a single trial prior to the receipt of reward, and during the interval between trials. It is characterized by instability, a relatively rapid rise during the interval of a single trial. The process need is assumed to be a decreasing monotonic function of delay of commerce with a goal object and a decreasing monotonic function of time following commerce with a goal object.

The process need is also assumed to increase or decrease in strength with variations in the external stimulus situation in the absence of consummatory activity or the attainment of a goal object.

It is hoped that this rather vague connotational analysis of the process need variable will delimit the meaning of the term sufficiently for further discussion. Its meaning perhaps will be more explicit following a survey of some relevant experiments. There are a number of such experiments which do not appear to be readily accounted for in terms of state variables which contribute to the reaction potential (sE_R) determining behavior.

II

A group of experiments bearing on the present problem are those employing some form of preliminary reward. According to the present formulations the introduction of a reward prior to an experimental trial arouses a process need which produces an increment in the reaction potential determining performance. Of course, where large amounts of reward are given prior to an experimental run a reduction in the tissue need state will result in a decrement in the level of the effective drive state. This would offset or prevent entirely the development of the process need.

Bruce (5) introduced a preliminary reward in a study of the gradient of locomotion in a straight runway. Control and experimental Ss were given one trial per day while thirsty. When the experimental Ss were given half their daily normal water intake in the starting section immediately before a run, there was a significant increase in running time. An increase in running time in the first half of the runway also occurred when 25 per cent of the normal water intake was received before the daily run. These results undoubtedly

reflect the reduction of the thirst need, and are in accord with other experiments employing large preliminary rewards as a means of reducing drive strength (23, 29). However, a significant acceleration did occur in the first half of the runway when the experimental Ss received $\frac{1}{2}$ c.c. of water prior to their run. This significant acceleration did not persist after six trials of predrinking with $\frac{1}{2}$ c.c. of water.

These results indicate that at some point between 25 per cent of the daily normal water intake and $\frac{1}{2}$ c.c. the increments in the process need offset the reduction in need state and are sufficiently great to facilitate performance. The decreased facilitation of performance with repeated preliminary reward trials is of some theoretical interest. It may indicate that the process need aroused by preliminary reward does not continue to contribute to reaction potential in the same manner after repeated elicitations. A phenomenon similar to inhibition of delay may develop (22, Chap. 6). With repeated preliminary reward trials, the process need gradient following the reward becomes suppressed. The process need is aroused only immediately before and after attainment of the goal.² The post- and ante-dating gradients become differentiated.

The hypothesis that the process need will decay as a function of time following commerce with a goal object is indicated by a preliminary reward experiment conducted by Morgan and Fields (21). Five minutes after cessation of preliminary feeding on varying amounts, Ss were run in a maze or straight run-

way. The results showed that prefeeding as little as five per cent of the daily ration produced an increase in running time and errors. In a second experiment, Ss receiving five per cent of their daily ration during preliminary feeding were placed in the starting section of the maze immediately after cessation of eating. Running time now tended to be shorter, in agreement with the results obtained by Bruce.

Another experiment employing varying amounts of preliminary reward has been conducted by Maltzman (19). Five different groups of hungry rats received 0, 100, 400, 1200, and 2400 mg. of food approximately 20 seconds before running in a straight runway. All groups received a reward of 1000 mg. at the end of a run. It was found that the 1200 mg. preliminary reward group ran significantly faster than the control group receiving no preliminary reward. The other groups were not significantly different in their performance from the control group. These data again indicate that there is an optimal amount of prefeeding. Smaller amounts do not elicit a process need great enough to facilitate performance. Larger amounts of prefeeding reduce the strength of the primary tissue need state.

It has been fairly well established that the growth of hunger and thirst does not follow the same functional course. Thirst appears to reach its maximum more rapidly and may be more intense than hunger (33). Evidence obtained by Bruce (3) indicates that the process need counterparts of hunger and thirst likewise follow different functional courses.

The process need elicited by 10 seconds of predrinking following 24 hours of water deprivation increased sufficiently to facilitate maze performance in terms of an error criterion. It continued to increase even after Ss were permitted unlimited predrinking follow-

² Evidence that may be interpreted as indicating the development of inhibition of delay of the anterior gradient may be seen in the shifts in speed of running on the initial segments of the runway on successive days in Hull's study of the goal gradient of locomotion (13).

ing 24 hours of deprivation. The process need elicited by prefeeding following 24 hours of deprivation did not increase sufficiently to facilitate performance until 20 seconds of prefeeding was permitted. It did not continue to facilitate performance with intervals of prefeeding larger than 40 seconds.

Interpretations of these experimental results, of course, are complicated by the effects of the reduction in the primary tissue need state as a consequence of receiving the preliminary rewards. Another complication is introduced by the unknown effects of running the same Ss under a variety of different preliminary reward conditions.

Anderson (1), in one of his experiments on externalization of drive, found a significant increment in performance following prefeeding in a multiple unit T-maze. Eight hungry non-rewarded Ss were given 36 daily trials in a six-unit maze. They were then given six daily trials with a five second prefeeding period in individual feeding cages immediately before each daily run. As in the previous control period, no reward was obtained at the conclusion of a run in the maze. Significantly fewer errors were recorded in this prefeeding period than in the previous control period. A six-day period then followed in which prefeeding was eliminated. A significant increase in errors over the previous prefeeding period was obtained. The second introduction of prefeeding again resulted in a significant decrease in errors.

The change in performance described in Anderson's experiment may be accounted for on the basis of two assumptions. First, the Ss learned something about the maze during the no-reward period. In Hull's terms, some habit strength had been built up in the absence of reduction of the primary hunger drive. Perhaps escape from the blind alleys or some less obvious factor

was the source of reinforcement. The drop in errors in the performance curve of these hungry Ss during the non-rewarded period of training (2) may, in part, indicate the acquisition of some habit strength.

The second assumption is that the preliminary feeding resulted in an increment in the process need which would multiply the previously established habit strength. There would result from this an increment in the effective reaction potential determining performance. When the preliminary feeding was withheld, the increment in the reaction potential contributed by the process need was withdrawn. A decrement in performance resulted.

Other experiments showing improvement in performance following preliminary reward have been conducted by Hull (11), and Bruce (4).

In addition to the preliminary reward type of study, there is another type of experiment supporting the hypothesis of process needs. These experiments further indicate that there is a class of needs that is a function of initial conditions other than, or in addition to, deprivation of a goal object.

For example, Yoshioka (37) found, accidentally, that wild rats preferred large sunflower seeds to small seeds when both were equally available. Subsequent experiments with albino rats under controlled conditions produced the same results (36). Individual Ss were presented with a mixture of large and small seeds and allowed to eat for five minutes. Significantly more large than small seeds were eaten. No difference was found in the selection of seeds when Ss were fed in the dark. It was also found that the large seeds took longer on the average to eat than the small seeds. That the preference for the larger seeds was not due to their greater primary reinforcement value is indicated by the finding that the meat

of 50 large seeds weighed exactly as much as the meat of 50 small seeds. The longer time to eat the large seeds apparently was spent in cracking and opening them. More work was required to open large than small seeds; therefore, a greater amount of work inhibition should have been built up to the opening of large seeds. Since the amount of reward was the same, Ss should have selected the small seeds. Initially they did not, as previously indicated. A preference for small seeds did, however, develop after repeated trials (35).

According to the present formulation, the Ss persisted for some time in selecting the large seeds because of the greater process need elicited by the visual difference in size. After repeated trials the differential amount of inhibitory potential which favored selecting the small seeds became greater than the effective reaction potential for selecting the large seeds. A shift in preference would occur as a result.

Results of a related experiment have been reported by Katz (15). It was found that the amount of consummatory activity of hens increased with the introduction of larger amounts of food. In other experiments hens were permitted to eat until cessation. The remaining food was removed and then replaced immediately. The hens immediately resumed eating.

These results do not appear to be readily explainable in terms of the state variables of primary hunger needs and habit strength. They can be interpreted in terms of the hypothesis of an increased process need induced by the sight of the larger pile of food or the re-introduction of food.

A third class of evidence from which a process need can be inferred is the effect upon performance of upward or downward shifts in the amount of reward. Under certain conditions it has

been found that Ss receiving a given quantity of reward and then shifted to larger amounts run faster, while Ss shifted to smaller amounts run slower, than Ss consistently run to these amounts.

However, there appears to be some inconsistency between the results of the experiments investigating this problem (7, 38). Zeaman (38) failed to obtain contrast effects with approximately fourfold and twelvefold shifts in amount of reward. Crespi (7) obtained such effects with roughly the same shifts in amount. With larger, fiftyfold, shifts, Zeaman did obtain a significant contrast effect when reward was shifted upwards, but did not obtain such an effect with downward shifts. Crespi obtained significant effects in both directions of reward shifts.

These differences may be due in part to the qualitatively different rewards used in the two experiments. Crespi used dog chow while Zeaman used a process cheese as the incentive. These different incentives may produce characteristically different parameters determining the shape of the hypothetical process need function.

In accounting for the results of his experiment, Crespi (6) assumed that incentives rewarding the Ss' performance produce an "emotional drive." Different amounts of incentive produce different strengths of this "emotional drive" or "eagerness" for the incentive. The amount of eagerness or anticipatory emotional tension elicited by a given incentive-amount depends upon repeated commerce with the incentive. Anticipation of a given reward and failure to obtain this incentive-amount produces an emotional reaction. If the incentive is greater than previously attained and therefore anticipated, the emotional drive is increased; elation occurs. If the incentive is less than anticipated, there is again an increase in the emo-

tional drive, but now the drive is akin to anger. Competing responses, signs of frustration, are elicited, and a decrement in performance results. The frustration resulting from the receipt of a small incentive is greater than that following absence of incentive. This would account for the inferior performance of the small incentive group as compared with the no-incentive group.

The interpretation of an "emotional drive" produced by the incentive is analogous to the process need construct suggested in this paper. It is therefore possible to translate Crespi's interpretation of the contrast effect into behavioristic terms. This task requires several additional assumptions. First, that there is another class of needs that has the properties of a process variable, frustration, which is a function of the antecedent condition of withdrawal of previously experienced reward. Unfortunately, at present there is even less experimental evidence for this construct than for the appetite process needs. The second assumption is that the amount of frustration induced is directly related to the strength of the process need present. Finally, the hypothetical process need function is ogival or S-shaped. A function of this kind representing the process need is probably unduly complex. It may represent the increments in the process need as well as decrements produced by need reduction following receipt of the incentive. It is possible that this assumption can be derived from other principles. At best it is a very crude first approximation. It is, however, in keeping with the results of Bruce (3) on prefeeding, and some testable implications can be drawn from even this rough approximation.

III

The studies briefly surveyed have been cited as evidence for the hypothesis that there is a class of needs that is

a function of initial conditions such as stimulation of distance receptors by a goal object, limited consummatory activity, and possibly, commerce with stimuli associated with consummatory activity. The entailed operations and consequent facilitation of designated criteria of performance constitute the basis for an operational definition of the process need.

Although the process need is a function of the stimulus situation exciting the organism at the moment, the extent to which the process need is dependent upon the antecedent condition of learning cannot be precisely determined at present because of the lack of experimental data. An observation by Pavlov is, however, relevant here (22, p. 22 f.). He reports that dogs raised on a diet free of hard food do not salivate at the sight of this incentive. Dogs raised on a normal diet do salivate at the sight of food. It is, therefore, reasonable to assume that the process need responsible for arousal of the anticipatory response of salivation is dependent upon past experience with food.

Several writers have theorized about the motivational condition that has been called a process need in the present paper. They have attempted in different ways to relate it to already existing constructs or to formulate new constructs to account for the behavior under consideration.

Spence (31) has suggested that the hypothetical fractional anticipatory response (rg-sg) may serve to motivate the organism as well as provide differential cues for responding. There is an increase in motivation with the elicitation of an anticipatory response by the original conditioned stimulus or generalized stimuli and the failure to immediately attain the goal object. Anticipation and non-attainment produce a conflict situation and a state of tension.

Spence (32) has also suggested an

alternative interpretation of the role of anticipatory responses in motivated activity. The hypothetical response-produced stimulus, sg , of the anticipatory response, rg , acts as a stimulus intensity dynamism, thereby producing an increased reaction potential when elicited. Furthermore, the anticipatory response is assumed to be a function of the amount of reward, not habit strength (31).

If the reduction of process needs to the stimulus intensity dynamism is possible, process needs would be interpreted as facilitating performance because of the increased intensity of stimulation following the initial conditions previously mentioned. It may also prove desirable to consider all classes of needs solely in terms of stimulus intensity (20). In this schema need states and process needs would be interpreted as facilitating performance because of the more intense stimulation following deprivation, or the additional increments in stimulation coming, perhaps, from the arousal of the fractional anticipatory response. A distinction between the two stimulus dynamisms would still be made because of the different operational definitions and functional relationships obtained in the case of the process need and the need state.

Another distinction similar to the one made here between need states and process needs has been offered by Guthrie (8). Following a criticism of his contiguity theory of learning by McCulloch (18), Guthrie made a distinction between annoyers, or primary tissue need states, and annoyances. The latter is produced when an S is placed in a particular problem situation. This, perhaps, arouses anticipation of success that cannot be immediately attained. Rewards reduce the annoyances temporarily, affording an immediate source of reinforcement. McCulloch points out that a pellet of food is not sufficient to

reduce the hunger drive. Learning occurs because the annoyance, or excitement induced by the immediate situation, is inhibited by the reward.

An alternative account, of course, would be that the source of immediate reinforcement is the secondary reinforcing property of the goal object which is independent of any change in process needs or need states.

Seward (26) has also considered incentives as a source of motivation. The surrogate goal reaction or anticipatory response elicited by an incentive or stimuli associated with it is assumed to be capable of activating habit structures in a manner similar to that of primary drives. In a more recent paper (25) he has hypothesized that the anticipatory response is a function of amount of reward and delay of reinforcement, and has assigned a fundamental "reinforcing," or perhaps more appropriately a "forcing," role to this mechanism. He has assumed that "reinforcement" produces increments in reaction potential not habit strength, and that "a response is reinforced when it is motivated by rs_G and rs_G is facilitated" (25, p. 370).

This review has been limited to appetite process needs that are a function of certain initial conditions. It may well be that the need conditions induced by frustration or the withdrawal of a goal object also have the functional characteristics of process needs (24, 27, 29). Similarly, performance determined by qualitatively different goal objects may be ascribed to variations in the parameters of the process need function. An analysis of these problems, however, is beyond the scope of the present paper. The hypothetical relations between the process need or its possible reduction to other principles or constructs in behavior theory also require experimental work and theoretical analysis that cannot be treated in this paper.

SUMMARY

The preliminary hypothesis presented was that in addition to primary and secondary drives, as they are ordinarily conceived, there is another kind of motivational condition determining behavior. This is the process need. It may be distinguished from other motivational constructs on the basis of its initial and consequent conditions and functional characteristics. Some of the relevant experimental evidence was surveyed as a basis for the assumption of such a construct.

REFERENCES

1. ANDERSON, E. E. The externalization of drive. IV. The effect of prefeeding on the maze performance of hungry rats. *J. comp. physiol. Psychol.*, 1941, 31, 349-352.
2. —. The externalization of drive. III. Maze learning by non-rewarded and by satiated rats. *J. gen. Psychol.*, 1941, 59, 397-426.
3. BRUCE, R. H. The effect of lessening the drive upon performance by white rats in a maze. *J. comp. Psychol.*, 1938, 25, 225-248.
4. —. The effect of varying the place of the fractional anticipatory consummatory response upon the rate of acquiring a simple learning problem. *J. gen. Psychol.*, 1943, 63, 165-175.
5. —. An experimental investigation of the thirst drive in rats with especial reference to the goal-gradient hypothesis. *J. gen. Psychol.*, 1937, 17, 49-60.
6. CRESPI, L. P. Amount of reinforcement and level of performance. *PSYCHOL. REV.*, 1944, 51, 341-357.
7. —. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, 55, 467-517.
8. GUTHRIE, E. R. The effect of outcome on learning. *PSYCHOL. REV.*, 1939, 46, 480-484.
9. HERON, W. T. Internal stimuli and learning. *J. comp. physiol. Psychol.*, 1949, 42, 486-492.
10. HULL, C. L. Behavior postulates and corollaries—1949. *PSYCHOL. REV.*, 1950, 57, 173-180.
11. —. Differential habituation to internal stimuli in the albino rat. *J. comp. Psychol.*, 1933, 16, 255-273.
12. —. *Principles of behavior*. New York: Appleton-Century, 1943.
13. —. The rat's speed of locomotion gradient in the approach to food. *J. comp. Psychol.*, 1934, 17, 393-422.
14. —. Stimulus intensity dynamism (V) and stimulus generalization. *PSYCHOL. REV.*, 1949, 50, 67-76.
15. KATZ, D. The vibratory sense and other lectures. *Univ. Maine Stud.*, 1930, 14, 1-163.
16. KENDLER, H. H. Drive interaction: I. Learning as a function of the simultaneous presence of the hunger and thirst drives. *J. exp. Psychol.*, 1945, 35, 96-109.
17. —. Drive interaction: II. Experimental analysis of the role of drive in learning theory. *J. exp. Psychol.*, 1945, 35, 188-198.
18. McCULLOCH, T. L. The role of clasping activity in adaptive behavior of the infant chimpanzee: III. The mechanism of reinforcement. *J. Psychol.*, 1939, 7, 305-316.
19. MALTZMAN, I. The effects of different amounts of preliminary reward on performance in a straight runway (in preparation).
20. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale University Press, 1941.
21. MORGAN, C. T., & FIELDS, P. E. The effect of variable preliminary feeding upon the rat's speed of locomotion. *J. comp. Psychol.*, 1938, 26, 331-348.
22. PAVLOV, I. P. *Conditioned reflexes*. London: Oxford Univ. Press, 1927.
23. REYNOLDS, B. The relationship between the strength of a habit and the degree of drive present during acquisition. *J. exp. Psychol.*, 1949, 39, 296-305.
24. ROHRER, J. H. A motivational state resulting from non-reward. *J. comp. physiol. Psychol.*, 1949, 42, 476-485.
25. SEWARD, J. P. Secondary reinforcement as tertiary motivation: a revision of Hull's revision. *PSYCHOL. REV.*, 1950, 57, 362-374.
26. —. A theoretical derivation of latent learning. *PSYCHOL. REV.*, 1947, 54, 83-98.
27. SHEFFIELD, V. F. Resistance to extinction as a function of the distribution of extinction trials. *J. exp. Psychol.*, 1950, 40, 305-313.
28. SIEGEL, P. S., & STEINBERG, M. Activity level as a function of hunger. *J. comp. physiol. Psychol.*, 1949, 42, 413-416.

29. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
30. SPENCE, K. W. The postulates and methods of behaviorism. *PSYCHOL. REV.*, 1948, 55, 67-78.
31. —. Theoretical interpretations of learning. Ch. 18 in *Handbook of experimental psychology*. New York: John Wiley & Sons, Inc., 1951.
32. —, BERGMANN, G., & LIPPITT, R. A study of simple learning under irrelevant motivational-reward conditions. *J. exp. Psychol.*, 1950, 40, 539-551.
33. WARDEN, C. J. *Animal motivation*. New York: Columbia Univ. Press, 1931.
34. WEBB, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, 39, 1-14.
35. YOSHIOKA, J. G. A further note on size preference of albino rats. *J. genet. Psychol.*, 1932, 41, 489-492.
36. —. Size preference of albino rats. *J. genet. Psychol.*, 1930, 37, 427-430.
37. —. Size preference of wild rats. *J. genet. Psychol.*, 1930, 37, 159-162.
38. ZEAMAN, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, 39, 466-483.

[MS. received December 6, 1950]

ON THE NATURE OF INHIBITION IN THE CEREBRAL CORTEX

BY MAX HAMILTON

University College, London

The celebrations last year of the centenary of the birth of Pavlov have reminded us that the study of the central nervous system, and especially of its higher functions, is a very young science whose origins are still within living memory. The great value of such celebrations is to make us pause for a while and attempt a new evaluation of the work of the pioneers in the light of further work in their subject. Such an examination gives us not only a renewed sense of perspective, but also a new understanding that indicates lines for further research.

This short essay arises from just such a careful examination of some of the experiments of Pavlov. It is concerned with a problem which is in a sense old, and which is yet with us: the controversy on the localization of cortical functioning. Many aspects of this controversy are now only of historical interest: The difficulties that arose from the comparison between different types of experiment and different species of animals have now been resolved, but others remain. Current controversies on the function of the motor cortex are concerned with one aspect of the nature of "local" and "general" functions of the cerebral cortex. Many of Pavlov's experiments cast a remarkably clear light on the nature of one of the most fundamental properties of the activity of the cerebral cortex: inhibition.¹ As far as it has been possible to check this, Pavlov

did not perceive the significance of his results, partly because of his bias in favor of the theory of the "functional mosaic," a theory of strict localization of the functions of the cerebral cortex. The number of experiments pertaining to this point is large, and their evidence is extraordinarily clear. I hope, therefore, that I shall be forgiven for dealing with what is essentially an experimental subject in a manner more suited to a literary investigation.

The first experiment I wish to quote is:

"One of the dogs had, besides different reflexes belonging to other analysers, a tactile conditioned reflex to acid, which had been experimentally generalized for the whole surface of the skin. The *gyri coronarius* and *ectosylvius anterior* were removed on the left side. On the fourth day after operation conditioned reflexes belonging to the analysers other than the tactile were present. The generalized conditioned tactile reflex returned on the eighth day, but only to stimulation on the left side of the animal, and soon reached its normal magnitude" (2, p. 346).

This means that the animal had a conditioned reflex to a tactile stimulus anywhere on its skin. It was operated on and the "projection" area of the skin removed on the left side of the brain, i.e., the projection of the right side of the surface of the body. After the animal had recovered from the immediate post-operative effects, it was found that it had lost the conditioned reflex to

¹ The cortical inhibition with which this paper is concerned is quite distinct from the function of the cortical "suppressor strips." These strips constitute part of the inhibitory

system dealing with motor activity, the lowest levels of which are concerned with such functions as reciprocal inhibition, and the highest with voluntary activity.

touch. This was restored after a few days, but only to the left side of the animal, corresponding with the untouched side of the brain.

In view of the intimate connection between the skin analyzer of one side and the analyzer of the other side, mediated by the *corpus callosum*, and demonstrated previously by the experiments of Anrep, it is not surprising that the uninjured skin analyzer (projection area) took longer to recover from the effects of the operation than did the rest of the cortex. Eventually, the tactile conditioned reflex returned on the right side as well, but meanwhile some remarkable results were obtained from the excised area:

"During the period while the conditioned stimuli of the above-mentioned places of the skin had lost their positive effect they divulged a definite inhibitory effect . . ." (2, p. 347).

That is to say, stimulation of the "refractory" area of skin (the right side) acted as an inhibitor to the positive conditioned stimuli, not only in the normal area of skin, but also in other "analyzers." Such an inhibition could become marked:

"Further, a repeated, and even more markedly a protracted, stimulation of these apparently ineffective places on the skin resulted in every experiment in a development of drowsiness and sleep, even in dogs which before the operation never showed any tendency to drowsiness in the stand. The sleep developed in these cases exclusively in connection with the tactile stimulation of these areas; apart from their application the dog kept completely alert during the experiment" (2, p. 347).

Stimulation of the skin whose cortical projection had been removed produced such intense and generalized inhibition that the animal was asleep. It is difficult to imagine how a nonexistent part of the cerebral cortex could exert any

influence on the rest, let alone such a powerful one. There is no doubt of the existence of some sort of activity related to the absent skin analyzer, as the next experiment shows:

"It was a problem of considerable interest to determine whether it would be possible by some means to disclose anything in the nature of a positive reaction to these tactile stimuli. We succeeded in demonstrating such an effect by the following modification of the experiments. The isolated stimulation of the refractory places of the skin was abbreviated from its usual duration of 30 seconds to 5 seconds. The abbreviated stimulus was now used several times in each experiment and at the end of the experiment, in order to test the reflex, reinforcement was again delayed for 30 seconds. Under such conditions it was possible to observe the positive as well as the inhibitory effect of the stimulation. The positive effect appeared quickly, *but was very small* (italics mine), and what is important, *disappeared while the conditioned stimulus was still acting*, whereas the effect of all other conditioned stimuli increased, as usual, towards the end of their isolated action as the moment of reinforcement approached." (A table is then given showing the details of the experiment.)

"A similar positive effect of stimulation of the usually ineffective places of the skin could be obtained also by some other devices—with the help of positive induction, by means of dis-inhibition, and by the use of caffeine" (2, p. 348).

This alleged activity of a nonexistent portion of cortex can be understood by assuming that not all of the skin analyzer of the cerebral cortex had been removed; or that adjacent parts of the cortex were able to partake of its functions to a slight extent—which comes down to much the same thing in the end. But the contrast between the feeble positive activity and the powerful inhibitory activity still requires explanation. It cannot be assumed that the power of inhibition in a given area

of cortex is greater than excitation, for in the first place there is no evidence for such a supposition, and in the second, the two processes are alike in many ways, e.g., in their rate of development, etc.

While it is possible to conceive that a small remnant of the cutaneous analyzer is able to exhibit feeble positive activity, it is obvious that something else must be involved in its inhibitory activity. Since the inhibition involves the rest of the cortex (as shown by the production of sleep), the phenomenon of inhibition involves at least a relation between the excised area of cortex and the rest of the cortex. Since the skin analyzer has been removed, it is a fairly obvious conclusion that the inhibition we observe in these experiments is really an activity of the remainder of the cortex. It might be suggested that these results are the products of the uninjured skin analyzer of the other side of the brain. This is disproved by the fact that normal uninjured cortex does not show such disparity between excitatory and inhibitory effects, and also by the fact that the uninjured skin analyzer behaves quite normally when it receives stimuli from its appropriate area of skin. We can interpret these experiments on the lines that excitation in the cortex is a function of a localized area, and inhibition is a function of the rest of the cortex, i.e., outside that area. The experiments separate the two functions by removing the localized area from which excitation arises.

As soon as we adopt the hypothesis that excitatory qualities of the cortex are the result of activity of a localized area, and that inhibitory qualities are the result of the activity of the rest of the cortex, a number of puzzling phenomena become perfectly clear. But before going on to these, it would be well to see if there is any evidence pointing directly towards this hypothe-

sis. The next experiment deals with this.

In the course of time the "refractory" area of skin regains its normal positive effect. By removing further portions of cortex, it was shown that this occurs as a result of "vicariation," i.e., the other areas of the cortex take over the specialized function of the missing area.

"... we now in some dogs extirpated the cortex completely on one side and studied the conditioned tactile reflexes from the skin of the opposite side of the body. . . . The results in all these experiments, in spite of many different modifications, were absolutely negative, and in spite of increasing the cortical excitability by strychnine and caffeine the cutaneous reflexes never returned. Experiments were also conducted to determine whether the tactile stimulation of the injured side of the body would exert any inhibitory influence upon other conditioned reflexes as it did after partial destruction of the cutaneous analyzer. *In the cases of complete unilateral extirpation of the cortex no such inhibitory influence of the tactile stimuli of the affected side upon the various positive conditioned reflexes (including tactile reflexes from the normal side of the animal) could be observed, whether as an after-effect or during the actual administration of the stimuli*" (italics mine) (2, p. 349).

Thus complete excision, and only complete excision, of the cortex of that side gives rise to the disappearance, not only of vicariation, but also of inhibition. It is scarcely possible to avoid the conclusion that the inhibition that is supposed to come from the extirpated area comes really from the rest of the cortex, and what is more, from the cortex acting as a whole. Incidentally, this result further disproves the suggestion that the opposite skin analyzer is the cause of the weak excitation and strong inhibition found in the first experiment.

It is probable that inhibition is always present in some degree. This is shown by an experiment in which part

of the cortex was excised. As a result, conditioned reflexes temporarily disappeared, and were then followed by a slow recovery.

"When the conditioned reflexes get re-established they are found not only to regain their normal strength but often to exceed it, often becoming considerably more stable than before. The inhibitory process, on the other hand, grows weaker. . . . Thus, after extirpation of a part of the acoustic area of H. Munk in two dogs, the conditioned alimentary reflexes not only made a complete recovery but considerably increased in strength, and kept constant throughout every experiment, whereas before the operation they used to decrease considerably towards the end of an experiment" (2, p. 324).

Removal of part of the cortex would appear to lead to a strengthening of the positive or excitatory process. A simpler and better explanation, as Pavlov himself recognized, is that removal of part of the cortex weakens inhibition. From this we may infer that inhibition is quantitatively related to the total area of cortex.

From this hypothesis we can deduce that excitation, which is a function of a localized area of cortex, should show general characteristics different from inhibition, which is regarded as an activity of the whole of the cortex. Inhibition should tend to show a more generalized character than excitation.

"Another peculiarity with respect to inhibition is observed after surgical interference. The inhibitory process becomes inert; and so to speak inflexible. As we saw before, in normal animals the inhibitory after-effect becomes with practice concentrated with regard to its duration as well as its extent. In the post-operative period this concentration proceeds extremely slowly and is imperfect. *This inertia of the inhibitory process is observed not only in the reflexes belonging to the analyser which has been surgically*

damaged, but also in reflexes belonging to other analysers" (italics mine) (2, p. 325).

Thus with the reduction in the quantity of cortex, the function of inhibition loses some of its delicate subtlety, but it does not become localized, it still retains its generalized character.

Differentiation of the stimuli to conditioned reflexes is essentially a process of inhibition. After the excision of part of the cerebral cortex:

"... the development of differentiations and conditioned inhibitions becomes more difficult and very often salivary secretion is observed in between the application of the stimuli—this never happening before" (2, p. 99).

This is another example of quantitative reduction, and it is confirmed by another experiment: this animal, after operation, developed convulsions. After these convulsions it ceased to move about and stood still for long periods. It also gave evidence of cutaneous hyperaesthesia.

"The post-mortem examination showed that the actual destruction involved in the main the posterior part of the cortex, the anterior part being affected to a much smaller extent. Naturally, therefore, the behaviour of this dog, as shown by observation of its conditioned reflexes, resembled that of animals in which the posterior part of the cerebral cortex had been extirpated. It is difficult, however, to find an adequate explanation for the prolonged standing of the dog in one place and the exaggerated reactions to touching the skin. So far as the former is concerned, it is impossible to decide whether it is an expression of a dominance of inhibition in the cortex subsequent to the periods of violent excitations (convulsions), or whether it should be regarded as a result of a partial damage to the cutaneous analyser. It is difficult to reconcile the extreme excitability of the cutaneous analyser with the hypothesis of a predominance of the inhibitory process in the cortex" (2, p. 375).

This last sentence clearly describes a situation which is in complete accord with the present hypothesis.

In an experiment designed to convert an inhibitory stimulus (a metronome beating at 60 per minute) into an excitatory stimulus, a number of anomalous results were obtained:

"To sum up, these experiments show that the transformation of an inhibitory point of the acoustic analyser into an excitatory one occurred only gradually and imperfectly. Moreover, and this is important, it rendered this point abnormal so that its stimulation by the conditioned stimulus of the metronome immediately led to a profound disturbance in the activity of the *entire* cortex, leading finally to an inability to withstand any strong conditioned stimuli without passing into different phases of inhibition, including the phase of complete inhibition" (2, p. 310).

Once again inhibition is seen to be, this time while undergoing disruption, a function that is related to the entire cortex, and to show generalized properties. It is particularly important that, in this experiment, we are dealing with a localized function, an auditory stimulus.

Except under certain specially limited experimental conditions, conditioned reflexes are generally differential conditioned reflexes depending upon conditioned inhibition. A "simple" conditioned reflex is not as simple as it seems. The whole of the environment of the

animal impinges upon it as a mass of stimuli, which has to be differentiated from the variable stimulus to which the animal is "learning" to respond. Thus differentiation and inhibition cannot be ignored even in the simplest of conditioned reflexes. Lashley has shown that a rat learns and relearns its way through a maze at a rate which depends on the area of cerebral cortex available to it. From this he deduced his theory that learning is a function of the cerebral cortex acting as a whole. The present hypothesis, that inhibition is a function of the cerebral cortex acting as a whole, thus forms a simple and clear link between the two types of experiment.

SUMMARY

An hypothesis on the nature of inhibition in the cerebral cortex is proposed, based on certain experiments of Pavlov. Certain deductions from it are made and found to be confirmed by other experiments. It is pointed out that this hypothesis links together very different approaches in this field, and even apparently opposed theories.

REFERENCES

1. LASHLEY, K. S. *Brain mechanisms and intelligence*. Chicago: Univ. of Chicago Press, 1929.
2. PAVLOV, I. P. *Conditioned reflexes* (trans. by G. V. Anrep). Oxford: Oxford Univ. Press, 1927.

[MS. received December 8, 1950]

RESPONSES IN ABSENCE OF THE ACQUISITION MOTIVE

BY WILSE B. WEBB

Washington University

In general, learning theorists have come to agree that responses in a particular situation are mediated both by habits (knowledges) and drives (motives). Diagrammatically, this relation may be given as follows:

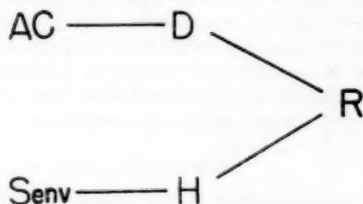


FIG. 1. Where AC = conditions leading to a drive state such as a deprivation schedule or a noxious stimulus; Senv = the environmental stimuli; D = the drive state; H = habit strength and R = the response which results from the joint effect of both drives and habits.

The continuation of a response in absence of the original motive under which the response was acquired has been a point of interest and of occasional controversy in behavior theory. Stated in less general terms, it has been frequently observed that organisms may learn a response under an original condition of motivation, but subsequently the response may occur when apparently the original motive no longer exists. Examples are numerous. A person learns to wear a tie under some urgings by his parents or some social pressure, but later he seems to wear the tie without the need of such urging. A person develops certain habits of frugality when he is working with limited funds. Years later, with unlimited funds, he is still pinching pennies from force of habit. The problem is obviously related to more extreme cases of continuing responses, such as fixations and phobias, etc.

In general, four types of theories have been developed to account for this behavior.

Type I. Nominal Theory. In absence of the drive of acquisition the response is elicited as an autonomous process.

Type II. Irrelevant Drive Theory. In absence of the drive of acquisition the response is elicited by the presence of other unreduced drives.

Type III. Acquired Tendency Theories. Stimuli, through association with an original drive, may come to evoke the response in absence of the original drive of acquisition. These theories have the same basic assumptions of acquisition. The theories may, however, be differentiated into two sub-theories in regard to the mode of evocation.

a. Acquired Drive Theories. In absence of the drive of acquisition the response is elicited by stimuli associated with an original drive, these stimuli having acquired the motivating properties of that drive and, as such, serving as evokers of the original response.

b. Stimulus Evoking Theories. In absence of the drive of acquisition the response is elicited by the presence of the original stimulus complex or components of this stimulus complex, these stimuli evoking the response as a function of the habit strengths developed during the original learning.

Type I: Nominal Theories

The first of these types of theories may be considered rather briefly since their development to date is rudimentary and the predictions evolving from such theories are limited. The author

has used the term "nominal theories" to describe those theories which name the event rather than specify the functional relations between the event in question and the hypothetical variable related to the event. As such, these hypotheses are analogous to the response-inferred theories described by Spence (19).

In one of the earliest considerations of the drive problem Woodworth noted that in many instances responses continued (occurred) in absence of apparent motivation (22). In accounting for this, Woodworth hypothesized that mechanisms, analogous to habits, acquired intrinsic drive properties within themselves. To quote Woodworth, "In short, the power of acquiring new mechanisms possessed by the human mind is at the same time a power of acquiring new drives; for every mechanism, at that stage of its development when it has reached a degree of effectiveness without having yet become entirely automatic, is itself a drive and capable of motivating activities that lie beyond its immediate scope" (22, p. 104).

In the development of the concept, Woodworth's writings closely approximate the Type III theory noted below and his hypothesis is labeled "nominal" only because of its relatively undefined character when referring to the independent variable of acquisition.

In 1937 Allport proposed the concept of "functional autonomy" to account for the behavior in question. This hypothesis merely stated that motives may become autonomous of their basic need states. Allport "... regards adult motives ... as self-sustaining, contemporary systems, growing out of antecedent systems, but functionally independent of them. ... Each motive has a definite point of origin which may lie in hypothetical instincts, or, more likely, in the organic tensions and diffuse irri-

tabilities. ... But as the individual matures the bond is broken. The tie is historical not functional" (1, p. 194).

Type II: Irrelevant Drive Theories

The second group of theories (Type II Theories), those dealing with the irrelevant drive phenomena, confine themselves to a very specific set of circumstances. Hull, in his formulation of his drive concept, suggested that the multiplicative relation holding between drive and habit may not only be a function of the dominant drive but also related "to certain residual amounts of various other drives" (7). These irrelevant or non-rewarded drives¹ were placed in an additive relation with the dominant drive. These irrelevant drives were further hypothesized to "have the capacity to sensitize habits not set up in conjunction with them." As such, it would be predicted from this hypothesis that in absence of the dominant drive, or here, the drive of acquisition, behavior would be evoked by these irrelevant drives. Hull used this set of hypotheses to explain the occurrence of responses in satiated animals which had been obtained in several previous experiments (23, 18).

Referring to Fig. 1, this theoretical formulation of Type II theories may be diagrammed as shown in Fig. 2.

It is obvious that because behavior does continue beyond the satiation of the relevant drive, such behavior cannot be attributed *a priori* to the role of irrelevant drives. Such behavior may be interpreted by any one of the types of hypotheses presently offered in this paper. Perhaps the most effective argu-

¹ These drives, the irrelevant drives, are defined as those drives present in the organism which have not been rewarded in the learning situation. Relevant drives or dominant drives, as described here, are those drives which have been reduced by this response, i.e., the original motive under which the response was learned.

ment for the attributions of behavior evocation to irrelevant drives would be found in the circumstances in which variations in these drives would lead

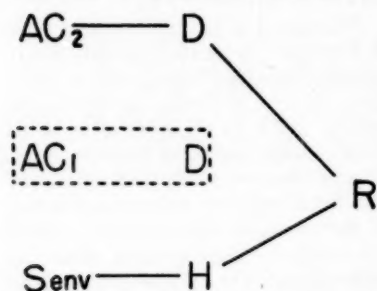


FIG. 2. When AC_2 indicates an irrelevant drive and AC_1 the original drive of acquisition which has been eliminated (indicated by the dotted line).

to a systematic variation of behavior evocation where secondary motivation, etc. would tend to be constant. Such evidence was obtained in an experiment by Webb in which systematic variation of the thirst drive (an irrelevant drive) led to a differential level of evocation of a habit pattern learned under the hunger drive (the drive of acquisition) (20).

There has been little discussion of the actual underlying mechanisms of these phenomena. Some consideration should, it is felt, be given the hypothesis that the operation of irrelevant drive may be described in terms of generalization along the drive continuum.

Type III Theories: Acquired Tendency Theories

Although the last hypotheses (Type III) have similar assumptions concerning the acquisition and evocation aspects, it seems possible that a critical difference is present in the assumptions concerning the "intervening" mechanism. As we shall note, this difference results in differential predictions which

are the critical differentia of any set of hypotheses.

Both theories assume basically that stimuli closely associated with a drive condition acquire, through this association, the property to evoke behavior in absence of the original motivation.

Type IIIA: Acquired Drive Theories

Most typical of the Type III theories are those of the first sub-category, i.e., acquired drive theories. These assume that the acquisition drive through some associative process is replaced by an acquired drive. The earliest of these theories to be placed in a systematic framework seems to be that of Anderson (2).² A simple statement of his theory may be quoted: "To summarize, it is assumed that a drive such as hunger is originally dependent for its arousal upon internal conditions of the organism but that, through continuous use of the drive in a relatively constant situation, the drive becomes aroused by an external situation in the absence of the original internal conditions, i.e., becomes externalized" (2, p. 224).

In the article in which this theory was derived, Anderson made some 29 predictions from this theory and in a series of experimental papers produced experimental evidence to confirm a number of these predictions (3, 4, 5). The majority of these predictions were correlated with the effect of the removal of rewards in the multiple maze situation and hence seemed directly related to the problem of secondary reinforcement (see below).

More recently the acquired drive theories have been considerably extended and systematized by the Yale group with the writings of Mowrer, Miller, Dollard, and Hull. Their theorizing

² May in an experimental article on acquired drives notes that Thorndike suggested a related point of view in the acquisition of wants and interests (9).

and experimental designs have been almost exclusively concerned with the acquired drive of fear or anxiety.

Mowrer's work in this area was introduced in 1939 in the form of two experiments (16, 17) and a theoretical article (15) which concerned themselves with the role of anxiety in the development of habits. Within this framework Mowrer was essentially more interested in the reinforcing properties of the anxiety reduction than with the acquisition and evocation properties of the anxiety drive in absence of the pain. Acquisition and evocation, however, are implicit throughout his writings.

Miller and Dollard (13) have presented the most complete theorizing on these acquired drives to date. These authors were particularly concerned about the underlying mechanisms through which these drives operate. In the book, *Social Learning and Imitation*, Miller and Dollard suggest "There is reason to believe that the drive value of self-induced stimuli is the basis of acquired drives . . .," with the drive capacity of these self-induced stimuli being based on a previous assumption that "a drive is a strong stimulus that compels action." In specifically referring to anxiety the process is described as follows: "Anxiety is one of the innate responses to pain. The physiological reactions producing the sensation of anxiety can easily be learned as responses to new situations, while those producing the original pain cannot. Therefore, anxiety is referred to as an acquirable drive and the pain as a primary drive" (13, p. 61).

In a more recent discussion of the problem the authors state, "We say fear is *learned* because it can be attached to previously neutral cues . . . ; we say it is a *drive* because it can motivate, and its reduction can reinforce, the learning and performance of new responses . . . (15, pp. 67-68).

The assumptions concerning the response-produced stimuli made by Miller and Dollard were discussed by May in an experimental paper (9). This author concluded that the essential response-produced stimuli were most likely "some pattern of neural activity in the motor areas of the central nervous system. . . ."

Miller has reported some three experimental papers investigating the drive quality of this acquired drive of anxiety (10, 11, 12). From these experiments he concluded that it was established that this acquired drive would exhibit the "functional properties characteristic of primary drives such as hunger . . ." (12).

All of the authors within the Yale group have implicitly, and on occasion explicitly (14), assumed that the principles holding for the acquisition of an anxiety drive also hold for other drive states. As such, their work is relevant to the problem in hand, i.e., in the evocation of behavior or responses in absence of the original drive of acquisition.

In passing, one should note the original attitude of Hull to these problems, for it reflects a prevalent problem in this area. Originally Hull tended to identify the problem of secondary motivation with secondary reinforcement. In Hull's terms, "The evocation of action in relation to secondary reinforcing stimuli or *incentives* will be called *secondary motivation*" (7, p. 226). Miller and Dollard, in their discussion, specifically defend a differentiation between the two problems (13). Hull, in a more recent statement concerning secondary motivation, states that "When neutral stimuli are repeatedly and consistently associated with the evocation of a primary or secondary drive and this drive undergoes an abrupt diminution, the hitherto neutral stimuli acquired the capacity to bring about the drive stimuli . . . which thereby become the condi-

tion . . . of a secondary drive or motivation" (8, p. 175). This latter statement seems to parallel, without the detailed specificity, the statements of Miller and Dollard.

The Type IIIA theories may be diagrammed in reference to Fig. 1 and Fig. 2 as follows:

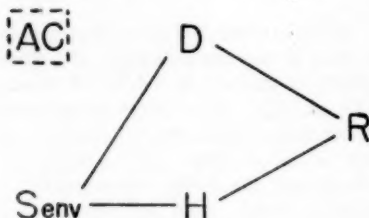


FIG. 3.

Finally, and most importantly, the implications for behavior theory of these acquired drive theories go far beyond the problem in hand. Functionally, these acquired drives are hypothesized to have the same dynamic characteristics as basic drives, i.e., may serve not only as energizers of already learned patterns of behavior *but* also as energizers of more general responses in new situations. In addition they have been hypothesized as entering into the development of new responses when the "drive states" are reduced. Their role as "drive stimuli" as cues to learning has not yet been designated.

In summary, considering the Type IIIA theories, there seems to be a general agreement that when neutral stimuli are closely associated with drive stimuli, these stimuli, through some associative process, may acquire the properties of the drive of acquisition. Behavior would be evoked in the presence of these previously neutral stimuli which are serving in a drive capacity.

Type IIIB Theories: Stimulus Evoking Theories

Although the Type IIIB theories have not been, so far as can be discov-

ered, systematized as such, it seems pertinent to discuss their possibilities. This set of theories would assume that stimuli which have been associated with the original behavior sequences come to evoke this behavior in absence of the original drive. This hypothesis would be differentiated from the ones discussed immediately above by the fact that no drive state would be invoked to account for evocation process in absence of the original drive, but would depend upon the habit strengths of the stimulus-response relations.

The Type IIIB theories may be diagrammed in reference to the previous figures as follows:

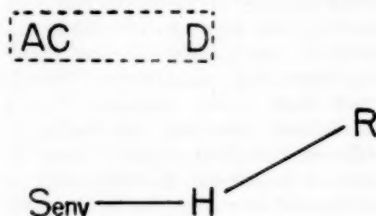


FIG. 4.

Although the author did not agree with such an interpretation, one form of such a hypothesis was stated by Williams in an early article (21). Given a situation in which a rat always found food in a white box, it would be possible that the following would happen: "The rat, whatever the environmental stimuli, responds with exploration . . . when he is experiencing the internal drive of hunger. . . . In the course of these explorations, he happens upon food . . . and eats it. . . . The sight of food becomes a substitute for hunger; that is a conditioned stimulus. . . ." From such a description one would only have to extend the statement to say that in absence of hunger the sight of food, acting as a conditioned stimulus, could evoke the original approach response on a pure habit

basis. Williams rejected that description but substituted what seemed to be a similar approach by calling the conditioned stimuli "signs," and described these signs by saying that "Their character consists behaviorally in the fact that their efficacy as stimuli depends on their function of preparing the organism for something that is to come. They initiate approach and avoidance responses, not in their own right, not even necessarily to or from themselves, but only because they have in the past been followed by certain other stimuli which are desired or feared" (21, pp. 493-494).

It is the first of these two descriptions that is particularly interesting when one considers the possibility that in the majority of the conditioning experiments there seems to be, at least implicitly, some such assumptions as noted in these Type IIIB theories. Initially it is necessary for a response to be evoked by some provoking stimulus (unconditioned stimuli). Subsequent to conditioning, however, these provoking stimuli no longer become necessary, and previously neutral stimuli acquire the property of evoking the behavior. When we recognize that drive in a broad sense may be described as an unconditioned stimulus, and we note the absence of the concept of drive in the conditioning design, it seems possible that many, if not most, of the classical conditioning experiments and a number of the instrumental conditioning experiments have implicitly assumed that merely the presence of previously neutral stimuli which have been associated with a response may now evoke this response in absence of any drive concept.

It will be recognized that this type of theory could be aptly described by the older "redintegration" theory of Hollingsworth (6). This theory suggested that where a part of the stimulus complex of an originally learned response

was present, "consequents may follow fragments or partial details of their former antecedents." If, then, the original external stimuli exist, the response may be evoked by a part of the stimulus complex in spite of the absence of the original drive. To quote Hollingsworth further, "... the motivated . . . sequences follow redintegration laws." A motive is defined by Hollingsworth as "... the former antecedent complexity as represented by a present partial detail." It may be assumed that the evocation is due to the original habit strength. Under such an assumption the redintegrative theory and the Type IIIB theory described here are identical.

This type of theory (stimulus-evoking) very nicely accounts for the common experience in which behavior which was originally "driven" or "voluntary" becomes automatic in character with little or no introspectively remaining drive in the behavior.

A final, but perhaps critical point, in regard to the Type III theories should be noted. It may be possible to derive the Type IIIA theory within the framework of the Type IIIB theory. Such a possibility may be most clearly seen if one focuses attention on the habit strengths of the internally produced responses in the original learning, an approach which has been followed by Miller and Dollard.

As stated by these authors, in regard to the acquired drive of fear, the argument would go as follows: "In short, we are assuming (1) that fear obeys the same laws as do external responses; and (2) that it has the same drive and cue properties as strong external stimuli . . . fear will be called a response . . . it will (also) be called stimulus-producing (14, p. 69). A similar statement of the same authors goes, "Fear is learnable in that it can be acquired as a response to previously neutral cues; the response to

those cues is a learned drive" (14, p. 88).

Their assumption then may be stated in terms of the acquired drive "stimulus" as "responses" which were originally conditioned to external stimuli. Agreeing that these "responses" are the "drive states" and that they represent conditioned or learned habit strengths the Type III theories become relatively indistinguishable.

This interpretation may be diagrammatically represented as:

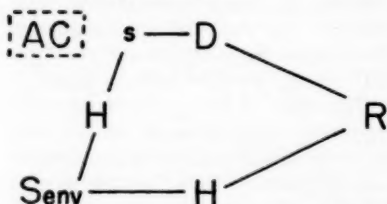


FIG. 5. Where *s* = internal response originally aroused by the AC conditions.

If we extend the Type IIIB theory in the direction of the Type IIIA theory as suggested above, the differences in prediction on the basis of the two theories become minimal. On the surface, the "acquired drive theory" is presently much broader and more flexible in describing behavior beyond that presently considered. This theory (the Type IIIA) allows the elicitation of new behavior patterns and further allows the learning of new behavior patterns when the original response is inadequate. Only by suggesting that the critical "habits" which are developed in the Type IIIB theory are those of continuing internal stimuli, does this latter theory allow such a prediction or description of behavior. If this latter is allowed, then the predictions of the two types of theories become identical and the problem remains as to the nature of the mediating structure, i.e., that of a drive or that of a set of habitual response-produced stimuli.

SUMMARY

We have reviewed various theories which have been proposed to explain the evocation of behavior in apparent absence of the original drive under which the behavior was acquired. One set of theories seems more descriptive than functionally useful. A second set, the irrelevant drive theories, seems experimentally verified but is restricted to a specified condition in which some alternative drive other than the original drive is definable. The third set of theories, those based upon assumptions concerning the acquisition of the evocation tendency as related to learning, seems the most widely applicable. The remaining problem for these theories seems primarily to center around the nature of what is acquired as the source for the behavior evocation.

REFERENCES

1. ALLPORT, G. W. *Personality*. New York: Henry Holt & Co., 1937.
2. ANDERSON, E. E. The externalization of drive: I. Theoretical considerations. *PSYCHOL. REV.*, 1941, 48, 204-224.
3. —. The externalization of drive: II. The effect of satiation and removal of reward at different stages in the learning process of the rat. *J. genet. Psychol.*, 1941, 59, 359-426.
4. —. The externalization of drive: III. Maze learning by non-rewarded and by satiated rats. *J. genet. Psychol.*, 1941, 59, 397-426.
5. —. The externalization of drives: IV. The effect of pre-feeding on the maze performance of hungry non-rewarded rats. *J. comp. Psychol.*, 1941, 31, 349-353.
6. HOLLINGSWORTH, H. L. *Psychology, its facts and principles*. New York: D. Appleton and Co., 1928.
7. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century Co., Inc., 1943.
8. —. Behavior postulates and corollaries. *PSYCHOL. REV.*, 1950, 57, 173-180.
9. MAY, M. A. Experimental acquired drives. *J. exp. Psychol.*, 1948, 38, 66-77.
10. MILLER, N. E. An experimental investigation of acquired drives. *Psychol. Bull.*, 1941, 38, 534-535.

11. —. The resistance to experimental extinction of anxiety as an acquired drive. *Amer. Psychol.*, 1946, 1, 290.
12. —. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, 38, 89-101.
13. —, & DOLLARD, J. *Social learning and imitation*. New York: Ronald Press, 1944.
14. —. *Personality and psychotherapy*. New York: McGraw-Hill Publications, 1950.
15. MOWRER, O. H. Anxiety-reduction and learning. *J. exp. Psychol.*, 1940, 27, 497-516.
16. —, & LAMOREAUX, R. R. Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.*, 1942, 54, No. 5. 34 pp.
17. —. Fears as an intervening variable in avoidance conditioning. *J. comp. Psychol.*, 1946, 39, 29-49.
18. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century Co., Inc., 1938.
19. SPENCE, K. W. The nature of theory construction in contemporary psychology. *Psychol. Rev.*, 1944, 47, 47-68.
20. WEBB, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, 39, 1-14.
21. WILLIAMS, K. A. The conditioned reflex and the sign function in learning. *Psychol. Rev.*, 1929, 36, 481-497.
22. WOODWORTH, R. S. *Dynamic psychology*. New York: Columbia Univ. Press, 1918.
23. ZENER, K. E., & MCCURDY, H. G. Analysis of motivational factors in conditioned behavior: I. The differential effect of changes in hunger upon conditioned, unconditioned, and spontaneous salivary secretion. *J. Psychol.*, 1939, 8, 321-350.

[MS. received December 18, 1950]

BEHAVIOR AND THE PSYCHOPHYSICAL METHODS: AN ANALYSIS OF SOME RECENT EXPERIMENTS¹

BY C. H. GRAHAM

Columbia University

A recent article (8; see also 9) presented an account of the way in which the topics of perception and the psychophysical methods relate to the study of behavior. The present discussion furthers the analysis by considering it in contexts provided by experimental programs in the field of perception. It is hoped that the treatment will have the advantage of making concrete the formal proposals of the earlier paper.

Three topics in the area of perception will be considered: (1) figural after-effects, (2) the influence of motives on discrimination, and (3) a theory of visual extent. All of these topics have provoked considerable research effort during the last several years, and they merit considerably more analysis than will be provided. In fact, the treatment will not present a detailed, critical account of the respective literatures. Rather, a small segment of experimental work will be considered against the background of a behavioristic account of the psychophysical experiment.

In order to clarify my position, I shall assert that the study of perception is an aspect of the study of behavior. This is an interesting statement, but it is of no consequence unless it can be demon-

strated that the data of perception experiments are specifiable in stimulus-response terms; that is, it must be shown that the data can be described by the relation

$$R = f(a, b, c \dots n \dots t \dots x, y, z). \quad (1)$$

This relation expresses the assumption of a behavioristic program: response is a function of certain specifiable variables. In particular, the first letters of the alphabet ($a, b, c \dots$) refer to properly specified aspects of stimuli; the last letters ($\dots x, y, z$), to properly specified conditions of the subject (physiological and inferred, including the effects of instruction stimuli); R , to response; n , to number of presentations; and t to time. The terms are not always independent of each other.

The discussion that follows will not be deterred by difficulties implicit in the concepts of *stimulus* and *response*. Nor will it be concerned with such important problems as the role of the instruction stimulus and the place of relatively unrestricted verbal behavior, that is, introspection. It will center about one general thesis: The types of stimulus-response functions that emerge from psychophysical experiments in the field of perception (and in particular, from the three areas of investigation under consideration) are special cases of the relation given in equation (1).

It will be obvious that, in the present context, the subject's verbal responses (e.g., "Yes" and "No") are considered to be samples of behavior. Such a viewpoint is not new. It is represented in

¹ This account was prepared under Project NR 142-404; Contract N6onr-271, Task Order IX, between Columbia University and the Office of Naval Research, U. S. Navy. Reproduction in whole or in part permitted for any purpose of the United States Government.

The article is a slightly revised version of a paper given at the Symposium on Conceptual Trends sponsored by the Division of General Psychology at the Annual Meeting of the American Psychological Association, September 5, 1950.

other discussions of psychophysics by Johnson (13) in 1929 and by Graham (7) in 1934; it may be implicit in the early article (1892) by Fullerton and Cattell (5). In Bergmann and Spence's nomenclature (1) the subject's verbal responses constitute an object-language that is to be contrasted with the experimenter's pragmatic metalanguage.

FIGURAL AFTER-EFFECTS

Consider first some problems of figural effects. The first systematic observations were made by Gibson (6), who found, among many other things, that if a subject regards a curved line during an initial inspection period, a subsequently presented straight line is said to be curved in a direction opposite to the curvature of the inspection line. More recently Köhler and Wallach (14) have conducted an extensive series of experiments dealing with figural after-effects; and they stress what appears to be an essential characteristic of the phenomena, that is to say, the displacement of the test object from the region of the previously viewed inspection figure.

To date, several studies have been concerned with the quantitative relations existing between the amount of displacement and certain controlling variables. At the moment one example will be considered in detail; that is, Hammer's experiment (10) on the time

course of figural after-effects. The precise treatment will deviate from the actual experiment in the sense that the method of constant stimuli will be substituted for the method of adjustment.

In Hammer's experiment the subject was instructed to regard a fixation spot in the field as shown in Fig. 1A. Below and to the left of the fixation spot is a vertical inspection line. The subject regards this line for a constant period of time, for example, a minute. At the end of the inspection time the inspection line is removed and the subject regards the fixation spot of the test field as shown in Fig. 1B. In the test field appears a vertical line at the same level but somewhat to the right of the original position of the inspection line. Above the fixation spot is a vertical comparison line which in successive test periods may be set in different positions on the horizontal axis.

According to Köhler and Wallach, pre-exposure to the inspection figure causes the subject to report that the test figure moves away from the position of the previous inspection figure. Operationally this means that when the subject is instructed to say, for various azimuth positions of the comparison figure, when the comparison and test figures are aligned, he will respond "Yes [they are aligned]" when the comparison figure has the position indicated in

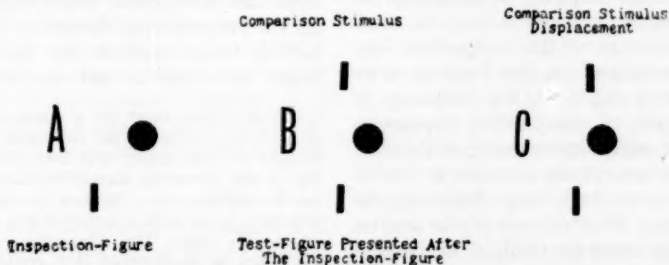


FIG. 1. Test- and inspection-figure configurations in Hammer's experiment (10). A. The inspection figure, regarded for 60 seconds. B. The test-figure (below) and the comparison-figure (above). C. Diagram of the comparison-figure displacement.

Fig. 1C. Due to pre-exposure to the inspection figure, the comparison and test figures are not aligned; and conventionally, but inaccurately, it may be said that the test figure has been displaced away from the inspection figure. In fact, it is the comparison figure that has been displaced.

What kind of psychophysical data are obtained in an experiment of the sort outlined? Assume (1) that the method of constant stimuli is used; (2) that the subject is instructed to respond in terms of alignment; and (3) that his responses are restricted to two categories: the positive, "Yes, they are aligned" and the negative "No, they are not aligned"; or more simply, "Yes" and "No."

Figure 2 presents a family of psychophysical curves that might be obtained in the experiment under consideration. Each curve refers to a given interval between the inspection period and the test period. The leftmost curve represents data obtained for a long inspection-to-test interval; the successively displaced curves represent data for shorter and shorter intervals. If we consider any single curve it may be observed that, for a constant inspection-to-test interval, the frequency of occurrence of the positive response is a sigmoid function of the position of the comparison line with reference to objective alignment. The frequency of occurrence of the positive response is low for small displacements of the comparison line. As the comparison line becomes more and more displaced, the frequency of occurrence of the positive response is low for small displacements of the comparison line. As the comparison line becomes more and more displaced, the frequency of occurrence of the positive response increases until, at fairly large displacements, it always occurs.

The stimulus-response function that is represented by each psychophysical

curve of Fig. 2 may be represented by the equation

$$R = f_1(a, t_1, x_1), \quad (2)$$

where a is the lateral displacement of the comparison line; t_1 is a constant value of inspection-test interval; and x_1 is made explicit as a measurable constant effect of instructions. Thus the

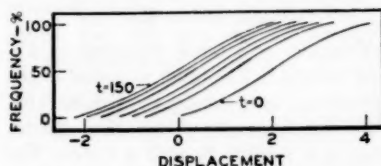


FIG. 2. A presumed family of psychophysical curves (method of constant stimuli) that might apply in Hammer's experiment (10). The rightmost curve refers to $t = 0$, where t is the time between the end of inspection and the appearance of the test-figure. The leftmost curve refers to $t = 150$.

data on figural after-effects given by the plot of a psychophysical function relate some measurable aspect of response to some measurable aspect of stimulus; and, in general, it may be said that a typical psychophysical function is, indeed, a stimulus-response relation.²

Only rarely in the study of perception are we interested in the psychophysical function itself. We are usually more interested in what may be called perceptual functions, i.e., functions that involve the determination of such a parameter as threshold. In obtaining these functions we use what might be called a null method that

² In the earlier paper (8) a certain amount of confusion attending the discussion of the method of limits might have been avoided if Fig. 3 had presented cumulative frequencies on the ordinate axis. In fact a cumulative frequency is the most meaningful way of presenting the data. The differential curve of Fig. 3 has the disadvantage that, except at the mean, the acceptance of a frequency measure of response change would lead to two values of threshold—an obvious absurdity.

determines how one stimulus variable varies as a function of another in order to produce a constant response effect.

Suppose that we regard Fig. 2 again, noting the positions of the successive curves. Each curve represents data obtained for a given time interval between the exposure to the inspection figure and the exposure to the test field. From each psychophysical function one can determine the value of comparison-figure displacement that corresponds to the 50 per cent occurrence of the positive response. When we determine this average displacement for each psychophysical curve and plot that displacement against the time interval between inspection and test, we obtain a perceptual function such as the one shown in Fig. 3.

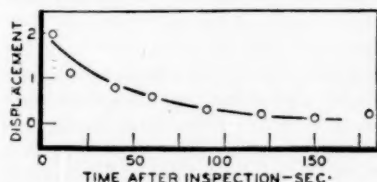


FIG. 3. The perceptual function indicating how figural-after-effect displacement varies with time between the end of the inspection period and the start of the test setting; i.e., it indicates the course of "disappearance" of figural after-effects, calculated from the data of Fig. 2. From Hammer (10).

This curve, in Hammer's terms, represents the course of disappearance of figural after-effects. It is a second order type of curve and shows the way in which a displacement, referable to a constant frequency of response, varies as a function of time interval between the inspection and test periods. Such a function is a *stimulus-stimulus* function for which response is constant; and it may be characterized by the relation

$$a_c = f_2(t, R_1, x_1), \quad (3)$$

in which a_c is the value of a correspond-

ing to the accepted criterion frequency of response occurrence, R_1 ; t is time; and the constant instruction-stimulus effect, x_1 , is made explicit.

Of course one might determine a family of perceptual curves. a_c is then a function of t and a parameter b (such as area), with R_1 and x_1 constants.

These, then, are some selected stimulus-response relations of a figural after-effect experiment. The treatment, as presented, is general and flexible; hence it has many advantages from a systematic point of view. It does not, however, afford a theoretical account of figural after-effects. Such an account must develop from a theoretically oriented, experimental program; a program that will seek to determine explicitly stated relationships among stimulus, response and whatever hypothesized variables become logically necessary. The hypothesized variables advocated by Köhler and Wallach in their satiation theory may turn out to be fruitful, but it is probably unprofitable at the present time to present an *a priori* judgment on the nature of theories that may advance our understanding of figural after-effects. A theory is a survivor of tests.

THE INFLUENCE OF MOTIVES ON DISCRIMINATION

In the last several years many experimenters have studied the influence of motives on perceptual functions; and the resulting data have been of interest to social psychologists, psychologists interested in personality theory, and experimental psychologists.

The recent rapid development of this subject matter has not been without its controversial aspects. The points at issue will not be discussed here; they have been well considered by others. Rather, two representative experiments will be considered to see how the program that they exemplify relates to a

systematic view of the field of perception. Without assuming a position for or against the technical correctness of the experiments, the discussion will be concerned with the broad question, What is the significance of the program for perceptual theory?

The experiment of Bruner and Goodman (3) may be taken as a first example. I am aware of the contradictory results obtained by Carter and Schooler (4) and Rosenthal and Levi (21; see also Pastore, 20); but I consider the experiment because of its implications, reserving judgment as to a final decision on the result. Bruner and Goodman used a standard psychophysical method; hence, their experiment does not confront us with embarrassing questions of response specification such as exist in researches where "verbal estimates" are considered to have quantitative implications.

The experiment of Bruner and Goodman shows the influence of a condition of the subject, called "economic deprivation," on a size discrimination. Groups of "poor" and "rich" children were instructed to set the area of a variable circular aperture equal to the area of a simultaneously present penny, nickel, dime, quarter or half dollar.

In matches made against cardboard discs the children showed only slight error in the adjustment of the diaphragm aperture; but with the coins an error of overestimation increased with coin denomination to a maximum for the twenty-five cent piece.

Figure 4 shows that when the data were broken down into scores for the "rich" and "poor" groups, the "poor" children gave greater overestimations than did the rich children.

Let us analyze the experiment. In this experiment, the subject, with appropriate verbal stimulation, manipulates a stimulus variable until it is discriminated as equal in area to a penny,

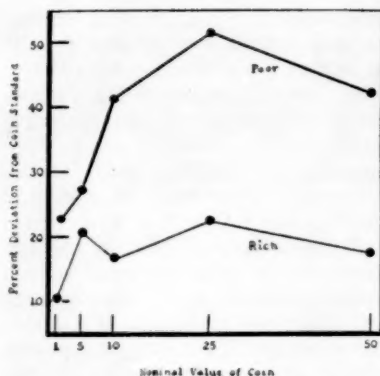


FIG. 4. Curves of size discrimination of coins for the "poor" and "rich" groups in Bruner and Goodman's experiment (3).

nickel, dime, quarter or half-dollar. The "rich" children, with a presumably constant low condition of "economic deprivation," give a curve, describing the relation between the comparison and standard stimulus, that rises from a low value at a low value of coin size and passes through a maximum at a medium value of coin size. The "poor" children, with a high degree of "economic deprivation," give a curve that is displaced upward on the comparison stimulus axis and has a higher maximum.

Thus the condition of "economic deprivation" is a determining parameter of the general size discrimination curve. Presumably, on a more finely graded scale of "economic deprivation," one could obtain a great number of curves. The general formulation that describes a finding of this sort may be represented as:

$$a_e = f_3(b, y, x_1, R_1), \quad (4)$$

where a_e is the "equality" setting of the comparison stimulus; b , the measured size of the standard stimulus coin or cardboard disc; y , a not exactly specified condition of the subject called "economic deprivation"; x_1 , a condition of

the subject determined by previous instruction; and R_1 , a constant response frequency.³

"The Bruner and Goodman study represents a specific approach to a general problem: How are discriminations influenced by different conditions of the organism? The correlated question of how the conditions may be specified is probably a more important one; it is certainly more important as soon as we progress beyond the stage of demonstrating that various conditions cause different discriminations. Without operational or theoretical specification of 'conditions,' it is unlikely that we can give a satisfying answer to the first question. While it may be true that specification of the organism's state as pain, set, need, anxiety, joy, etc., may, at the conversational level, establish the existence of a problem, we hope to develop behavior

³ The earlier reference (8) discusses the method of adjustment in considerable detail. The method provides only one response category, i.e., adjustment. "However, classification of response may be accomplished arbitrarily by the following device. Responses that lead to stimulus settings lying within plus or minus one standard deviation (or any other appropriate statistical fiducial limit) of the mean of the settings may be classed in the category 'Equal.' Responses that give settings more than a standard deviation below the mean are classed in the category 'Less.' Responses that give settings more than a standard deviation below the mean are classed in the category 'Greater.'" The average setting a_e of the distribution of readings may be taken to be that setting that represents the category "Equal."

In using this arbitrary device the measure of the response "Equal" may be taken as the frequency of occurrence of settings between the limits of plus and minus one standard deviation. The maximum frequency (i.e., the frequency at the average) suggests itself as the obvious acceptable measure R_1 of the response class "Equal." This treatment, of course, depends for its validity upon the relation that exists between the settings given by this method and those given by the other methods. Fortunately, the correlations are high; "equality" settings by the methods of constant stimuli and adjustment are comparable, as are the variability measure of the latter and the just discriminable difference of the former.

theory to the point where these unknown conditions may be described in such form that their quantitative influences on behavior will be predictable; and we must relate them to the inferred but quantitative variables that connect stimulus and response. One of our first duties is probably to see that definitions of the various 'conditions' are anchored to definable antecedents and consequents" (9).

The experiment by Lambert, Solomon and Watson (15) is of considerable interest in the present context, for it was performed with a view to specification of a condition called "subject value." The authors demonstrate that the establishment of a token-reward sequence results in the relative overestimation of token size by children, and that extinction of the sequence removes the overestimation tendency to a great extent.

The experiment shows that size estimations of children are influenced by the number of rewards and number of reward withdrawals. In general, it is shown that:

$$a_e = f_1(b, n_r, n_w, x_1, R_1), \quad (5)$$

where the letters a_e , b , x_1 and R_1 have the meanings previously given and n_r and n_w represent number of rewards and number of reward withdrawals.

The discussion implies that there is little need to argue concerning the problems of conditions of the subject as they influence behavior; nor is there need to re-examine the old problems relating to the empiricism-nativism controversy in perception. What is required is first, that "conditions of the subject" be specified adequately and secondly, that an improved experimental program determine the controlling parameters. The two programs must evolve together; and the question as to program then becomes: Does the experiment show that, in fact and with respect to a given example of behavior, a certain measurable

and specifiable variable has a value different from zero?

A THEORY OF VISUAL EXTENT

In 1947 the late Rudolph Luneburg presented a mathematical analysis of binocular vision (16). Luneburg was dissatisfied with certain exceptions to the classical laws of binocular space perception, and his theory is concerned with the problem of visual extent (16, 17, 18).

An example of an experiment on visual extent is the following: The subject is instructed to equate the distance between two variable comparison points until it equals the distance between two standard points. The resulting setting varies as a function of several controlling conditions, among which are the relative distances and angular relations of the comparison and standard points.

Specific problems that fall into the general class of experiments on visual extent are the horopter problem (see, for example, 9), the Hillebrand alley problem (2, 11), size constancy (9) and distorted rooms. Luneburg considers all of these phenomena as basic to his theory; but, at the moment, only the alley problem will be described.

The Hillebrand alley problem may be described in the following way: The subject is instructed to arrange two rows of light points in such a way that they seem to form parallel alleys symmetrical to the median plane. The settings made by the subject in terms of these instructions form diverging curves that are not physically straight. This type of finding is clearly not accounted for by a classical theory of corresponding points.

In order to present a systemization of the phenomena of visual extent Luneburg presents what he calls the mathematical form of the metric of the visual space, which he characterizes as a metric in a manifold of sensations.

The treatment expresses the general hypothesis that, for an individual observer, the apparent size of a line element is uniquely determined by its physical coordinates. Thus the discrimination of visual size depends not only upon angular relations but also upon the location of the line element.

In essence the theory says that a number, D , can be expressed as a complicated function of (1) the angles of elevation of the two points that limit a given extent; (2) the convergence angles of the lines of sight to the points; and (3) the visual angle subtended in each eye by the horopter distance between the two points. D can be computed for the comparison points and for the variable points; and, if the theory is correct, the two D 's should be equal when instructions are given to equate visual extents, regardless of the relative distances and angular relations of the two extents.

It is of considerable interest that the value D determined by these operations mathematically represents a distance projection in non-Euclidean space. For the psychologist this is an interesting but not final conclusion. The mathematical formulation may be of considerable value in systematizing many types of observation if it is found that the description does predict. But we are not likely to forget that Luneburg's symbolic system describes how a measured setting made by a subject varies as a function of a magnitude of a standard stimulus; and we are faced again with a relation of the basic type

$$a_e = f_s(b, R_1, x_1), \quad (6)$$

where a_e is the "equality" distance setting; b , the comparison distance; R_1 , the reference response frequency; and x_1 , the constant effect of instructions.

The psychologist will be interested to know if Luneburg's description will, in fact, predict; but he will continue to

seek an account that, in appropriate form, will not only predict but provide more complete analysis. Luneburg's mathematical account is an interesting intervening variable. It is not an hypothetical construct (19).

DISCUSSION

An attempt has been made to demonstrate that the data of certain experiments in the field of perception are in a form that accords with equation (1). Specific variations of equation (1) describe the stimulus-response relations of the various experiments. It is proposed as a general principle that an experimental function obtained with a procedure that uses a psychophysical method, involving appropriate response restriction and specifiable stimulus variables, is a special case of equation (1).

It may be conceded that the present analysis is interesting as a stimulus-response formulation of psychophysical experiments, but it may be argued that it does not provide a description of perception. Certainly it does not describe *perception* if by that term is meant any area of knowledge that exists between the limits of a consideration of sensory nerve impulses on the one hand and the language responses ("private experiences") of a clinic patient on the other. Such a usage, of course, deprives the term of any significance or specificity. The term might advantageously be restricted to the kinds of stimulus-response relations that are discussed in the present paper. Names appropriate to other kinds of experimentation might be applied as the bases for classification become better specified.⁴

In terms of the present formulation,

⁴ The recognition of a need for proper naming is not new. For example, "perceptual" experiments that involve verbal estimates have often been classified in a category of their own, i.e., as experiments in *judgment*.

It is also of some interest that the problem

any presumed differences between "sensory" and "perceptual" research evaporate. Both types of research give rise to the same sort of stimulus-behavior functions. Differentiation between the two fields has often rested, in the recent past, upon theoretical propositions: in the case of "sensory" psychology, statements couched in a "physiological" vocabulary; in the case of "perception," statements couched in a "psychological" vocabulary. Since the aim of the scientist working in this area is to make explicit the relations of equation (1) by whatever path his analysis may take, then the differentiation between "sensory" as "physiological" and "perceptual" as "psychological" is probably not definitive.

SUMMARY

Stimulus-response formulations of some experimental data in the field of perception are presented. The formulations are concerned with experiments on (1) figural after-effects, (2) the effect of motives on discrimination, and (3) a theory of visual extent. Specific variations of a general stimulus-response equation, equation (1), describe typical functions obtained in the various experiments. The term *perception* might be advantageously restricted to functions obtained with those psychophysical methods that use response restriction and specifiable stimulus variables.

REFERENCES

1. BERGMANN, G., & SPENCE, K. W. The logic of psychophysical measurement. *Psychol. Rev.*, 1944, **51**, 1-24.
2. BLUMENFELD, W. Untersuchungen über die scheinbare Grösse in Sehraume. *Z. Psychol.*, 1913, **65**, 241-404.

of psychological scaling reduces, in the present context, to an analysis of the psychophysical and perceptual functions that arise in the scaling operations. The primary interest lies in the analysis of the functions, not in their use as scales.

3. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1947, 42, 33-44.
4. CARTER, L. F., & SCHOOLER, K. Value, need, and other factors in perception. *Psychol. Rev.*, 1949, 56, 200-207.
5. FULLERTON, G. S., & CATTELL, J. McK. *On the perception of small differences*. Philadelphia: University of Penna. Press, 1892.
6. GIBSON, J. J. Adaptation, after-effect and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, 16, 1-31.
7. GRAHAM, C. H. Psychophysics and behavior. *J. gen. Psychol.*, 1934, 10, 299-310.
8. —. Behavior, perception and the psychophysical methods. *Psychol. Rev.*, 1950, 57, 108-120.
9. —. Visual perception. In Stevens, S. S. (ed.), *A handbook of experimental psychology*. New York: John Wiley, 1951.
10. HAMMER, ELAINE R. Temporal factors in figural after-effects. *Amer. J. Psychol.*, 1949, 62, 337-354.
11. HILLEBRAND, F. Theorie der scheinbaren Grösse bei binokularem Sehen. *Denkschr. d. k. Akad. d. Wissensch. Math.-nat. wiss. Classe*, 1902, 72, 255-307.
12. JOHNSON, H. M. Some fallacies underlying the use of psychological tests. *Psychol. Rev.*, 1928, 35, 328-337.
13. —. Did Fechner measure introspectual sensation? *Psychol. Rev.*, 1929, 36, 257-284.
14. KÖHLER, W., & WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Philos. Soc.*, 1944, 88, 269-357.
15. LAMBERT, W. W., SOLOMON, R. L., & WATSON, P. D. Reinforcement and extinction as factors in size estimation. *J. exp. Psychol.*, 1949, 39, 637-641.
16. LUNEBURG, R. K. *Mathematical analysis of binocular vision*. Princeton: Princeton University Press, 1947.
17. —. Metric methods in binocular visual perception. *Studies and Essays. Courant Anniversary Volume*. New York: Interscience Publishers, Inc., 1948.
18. LUNEBURG, K. The metric of binocular visual space. *J. opt. Soc. Amer.*, 1950, 57, 108-120.
19. MACCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, 55, 95-107.
20. PASTORE, N. Need as a determinant of perception. *J. Psychol.*, 1949, 28, 457-475.
21. ROSENTHAL, B., & LEVI, J. Value, need, and attitude toward money as determinants of perception. *Amer. Psychol.*, 1950, 5, 313.

[MS. received December 14, 1950]

THE INTERRUPTED TASK METHOD IN STUDIES OF SELECTIVE RECALL: A REEVALUATION OF SOME RECENT EXPERIMENTS

BY THELMA G. ALPER

Clark University

Since its introduction in 1927 by Zeigarnik (24), the interrupted task method has been widely accepted by psychologists for the study of selective recall.¹ Recently, however, Sears (22, pp. 112-114) and Cameron (22, p. 194) have challenged its usefulness for this research area. They are critical largely because the method appears to have yielded diverse results in three seemingly similar experiments: Rosenzweig's (17), Alper's (1) and Glixman's (10). In evaluating these experiments, the reviewers relied heavily on Glixman's (9) discussions of the three studies, accepting as valid Glixman's two major assumptions: (1) The "three experiments were designed in much the same manner to answer the same question" (9, p. 491); and (2) Rosenzweig and Alper used "an inadequate measure of recall" once they had obtained their raw data. Since the three experiments *do* differ in orientation and design, acceptance of these assumptions introduces new confusions into selective recall research. The measure of recall which Glixman would substitute obscures the important theoretical differences underlying the three designs. Glixman's position, therefore, must be re-examined.

Glixman's argument runs as follows. Since Rosenzweig and Alper each included a stressful, as well as a non-stressful, condition in their experimental designs, they, like Glixman, must

have been interested in the effects of different degrees of experimentally induced stress on the recall of incompleting tasks and on the recall of completed tasks. To measure these effects, separate intersession comparisons for the recall of incompleting tasks and for the recall of completed tasks are required. Yet, Rosenzweig and Alper used an intrasession measure, the number of completed tasks recalled in a given experimental session being compared with the number of incompleting tasks recalled within that *same* session. Glixman's (9, p. 492) objection to the intrasession measure is that it cannot adequately reveal the effect of stress on selective recall since "any change in the recall of incompleting activities is a function not only of the increase in stress, but it is also a function of the change in the recall of completed activities." An intersession score, he says, would separate the two factors; an intrasession score would confound them.

In order to rectify Rosenzweig's and Alper's "error," Glixman reanalyzed their original data in terms of his intersession measure. These reanalyzed data yield results that differ both from the original intrasession findings reported by Rosenzweig and Alper, and from Glixman's own results. The three experiments now seem to contradict each other. Glixman, however, attempts to fit them all into a single theoretical framework (10). At this point the reviewers are unwilling to follow Glixman. They stress the diversity of results, not Glixman's theoretical reconciliation. In-

¹For summaries of the literature see Alper (1), Barret (6), Prentice (15), Rapaport (16) and Sears (21).

deed, the complexity of the reconciliation seems only to convince them that the basic difficulty must lie with the interrupted task technique itself. They recommend, therefore, that future experiments in selective recall be differently designed.

The reader who is only superficially familiar with the original papers under discussion would find the reviewers' treatments convincing. He would, accordingly, conclude that: (1) Glixman is correct in rejecting Rosenzweig's and Alper's use of an intrasession score; (2) selective recall is unlawful, or, at the least, these three experiments fail to reveal its lawfulness; and (3) the interrupted task technique is an unsatisfactory method for studying selective recall. A re-evaluation of the experiments in question will show, however, that none of these conclusions is warranted.

In the present paper the hypotheses, methodology and findings of the Rosenzweig, Alper, and Glixman studies will be re-examined. It will be shown how the three experiments differ in experimental design, and why an intersession recall measure would be incorrect even in Glixman's experiment. The point will be made that just as an experimental hypothesis determines the design of the experiment, so, too, the experimental method sets limits on the way the data are to be treated.

The rationale of the three studies can be understood best in historical perspective. Therefore, Zeigarnik's (24) original experiment and an early study by Rosenzweig and Mason (18) will also be considered here.

Working within the general framework of Lewin's (13) tension theory, Zeigarnik designed an experiment to measure the relative effectiveness for recall of incompleteness as compared with completed tasks. The experimental tasks were heterogeneous, some easy (e.g., stringing beads), others more diffi-

cult (e.g., solving jigsaw puzzles), the experimental setting informal and, presumably, non-ego-threatening. The interruption of half of the tasks was accomplished in such a manner that *S* was not aware that interruption *qua* interruption was methodologically important. The experimental session was in no sense a test of *S*. This was made clear to *S* in the instructions and in the atmosphere of the sessions.

Zeigarnik's recall measure was the ratio of the number of incompleteness tasks recalled to the number of completed tasks recalled. This ratio, known now as "the Zeigarnik ratio," is, of course, an intrasession measure of recall. If *S* recalls an equal number of both types of tasks, the value of the ratio, *P*, is 1. If more incompleteness tasks are recalled, *P* is greater than 1; if more completed tasks are recalled, *P* is less than 1. Zeigarnik found that the average *P* value for a group of college students was 1.9. For a group of adolescent children, aged thirteen and fourteen, the average *P* value was 2.1. In other words, most *Ss* recalled a preponderance of incompleteness tasks. In later studies, this datum has been referred to as "the Zeigarnik effect." As all later selective recall investigations have also revealed, some *Ss* recalled more completed than incompleteness tasks. In Zeigarnik's experiment these *Ss* were in the minority.

In explanation of the superior recall of incompleteness tasks, Zeigarnik postulated that when *S* accepts *E*'s instruction to work on a given task, a tension is set up within the *S* for that task. Completion of the task tends to resolve the tension system connected with it; interruption of the task tends to prevent the normal resolution of the tension. She further postulated that *E*'s question, "What have you been doing this hour?" also sets up a tension. This latter tension is resolved by recalling the tasks. Thus, at the moment of recall,

two vectors deriving from these tensions are present: one directed towards the completion of incomplete tasks, the other towards the recalling of tasks. The former, according to Zeigarnik, becomes effective for recall also, since recall as well as resumption of the incomplete task can relieve the unresolved tension. Traces of incomplete tasks would thus be more available for recall at the end of the experimental session than traces of completed tasks, since the two vectors favor the recall of incomplete tasks and only one favors the recall of completed tasks. While Zeigarnik did not test the statistical significance of the obtained differences, the superiority of incomplete tasks for recall held up in a variety of supporting experiments.² Zeigarnik's theory does not cover reversed recall ratios, i.e., the superior recall of completed tasks. Nor, indeed, would it cover an equal recall of both types of tasks.

The next important step, methodologically, was taken by Rosenzweig and Mason (18) who utilized the interrupted task technique in an effort to test Freudian repression experimentally. Using a group of crippled children whose ages ranged from 5 years 6 months to 14 years 8 months, these investigators intentionally employed a less impersonal setting than did Zeigarnik. The authors state (18, p. 249): "To arouse a genuine need in the subjects, the experiment was conducted as a contest. The children were informed by one of their teachers that they were to be given a test to determine how well they could do puzzles and that a prize would be awarded to the one who did the best." The Ss were tested individually.

The "genuine need" which the Es intended to arouse in the Ss was, of course, ego-involvement. In Zeigarnik's experi-

ment, the Ss were supposedly merely task-involved.

Elsewhere, Alper (4) has drawn the distinction between task-orientation and task-involvement, and between ego-orientation and ego-involvement. This distinction points up the difference between what the E intends by virtue of his instructions (orientation) and what S experiences (involvement). Presumably, involvement coincides with orientation. But it does not inevitably do so, as will be indicated later.

In Zeigarnik's experiment, the instructions were task-oriented. If the Ss were merely task-involved, incompleteness would be experienced as experimenter-induced and completion merely as task-compliance on their part. In the Rosenzweig and Mason study, the instructions were ego-oriented. If the Ss actually were ego-involved, incompleteness could well be experienced as personal failure, completion as personal success. Provided the Ss actually were ego-involved, the introduction of the ego-threat, essential to an experimental test of Freudian repression,³ would drastically alter the original conditions as utilized by Zeigarnik. This, at least, was the intention of Rosenzweig and Mason. We can now turn to the analysis of their data.

Since not all Ss were given the same number of tasks, Rosenzweig and Mason substituted for the Zeigarnik recall ratio an intrasession recall difference score computed as follows:

$$\frac{C's \text{ recalled}}{C's \text{ given}} - \frac{X's \text{ recalled}}{X's \text{ given}} = R. D.$$

³ While the writer would question whether failure on half of a set of jigsaw puzzles would, for the average child, be sufficiently threatening to ego-status to constitute the basic conditions for "Freudian repression," this need not concern us here. The important point is that the conditions in the Rosenzweig and Mason experiment were objectively different from the conditions in Zeigarnik's experiment.

² See Koffka (12, pp. 334-341) for an English summary of Zeigarnik's (24) experiments.

In this formula C stands for completed tasks, X for incompleting tasks, and R. D. for recall difference score. The R. D. is, of course, a percentage score.

Several interesting findings appeared in this study. Of major significance for the present discussion is the fact that no general recall tendency, such as Zeigarnik had found, emerged. The 40 children who served as Ss fell into four recall groups. Sixteen children recalled more completed than incompleting tasks; thirteen, more incompleting tasks; nine, an equal number of both types; and two children recalled none of the tasks. The authors attempt to explain only the first two recall patterns.

Each child was rated by his teachers on the personality variable "pride."⁴ These ratings were then correlated with the recall patterns. Children who were high on pride tended to recall a preponderance of completed tasks. These children also tended to have the higher mental age scores. Children who received the lower pride ratings recalled either a preponderance of incompleting tasks or an equal number of completed and incompleting tasks. The low pride children tended to have the lower mental age scores.

While the authors recognize that "the reliability of the results on pride is not very great" (18, footnote 1, p. 258), they nevertheless interpret their findings as follows (18, p. 258): "It would seem to be that, given an individual of sufficient intellectual maturity and a commensurate measure of pride, experiences that are unpleasant because they wound self-respect—perhaps it should be added in a social situation—are, other things being equal, less apt to be remembered than are experiences that are gratifying to the ego. This is in keeping with the

Freudian theory of repression." To account for the recall pattern of the low pride group, the authors assume that these children did not experience the situation as threatening. Rosenzweig and Mason (18, p. 259) state that these children "were presumably not enough wounded for repression to operate. Their reaction was, however, very definite and consistent. It is as if they took a more objective attitude toward the test, regarding the puzzles as so many external obstacles to be overcome. Their need for mastery of the environment was aroused but failure to satisfy it did not wound their egos or arouse feelings of inferiority. If anything, feelings of protest and perseverance were awakened by failure and these brought about a preference for X's [incompleting tasks] in recall. The fact that the unsuccessful puzzles were *unfinished* may also have created a perseverative tendency, such as probably prevailed in the experiment of Zeigarnik, making for the recall of X's. Such a perseverative tendency must have been present in the case of the positive [superior recall of completed tasks] group, too, but there the competing repressive tendency was so strong that the perseverative one was overpowered. In the case of the negative [superior recall of incompleting tasks] group the attitude of protest cooperated, rather than competed, with the perseverative tendency, and so the X's were better recalled." The authors recognize, thus, that involvement does *not* necessarily coincide with orientation.

Although the theory, as outlined, is highly suggestive, it lacks internal consistency. In the above quotation the authors seem to be postulating three separate tendencies: perseveration, protest and repression. Yet, in the theory, they deal only with perseveration and repression. Indeed, a "feelings of protest" concept in the low pride group, presents certain difficulties. For ex-

⁴ "Pride" was defined by these authors (18, p. 257) as "desire to stand well with the group, and pleasure in one's own achievement."

ample, it is not at all clear why "immature," "low pride" Ss whose "egos were not wounded" should experience in this situation "protest," or, more especially, "failure."⁵

Tentative as the theory is, the speculations nevertheless carry us considerably beyond Zeigarnik's experiments, since the groundwork is now laid for the possibility that factors other than the objective conditions set up by *E* may influence the direction of selective recall. Zeigarnik (24) had already noted in her experiment that particularly "ambitious" Ss were likely to show even more marked tendencies than "nor-

mal" Ss to recall incompleting tasks.⁶ Now, in the objectively ego-threatening situation set up by Rosenzweig and Mason, "proud," "mature" Ss tended to recall more completed tasks, while the "low pride," "immature" Ss who were presumably not ego-involved tended to recall more incomplete tasks.

In his subsequent research Rosenzweig did not pursue further the role of personality factors in selective recall. Instead, he designed the experiment (17) which Glaxman (9) contends was incorrectly analyzed. Taking as his cue the differences in recall of the "non-threatened" vs. the "threatened" children, in the new experiment, Rosenzweig intended to distinguish more sharply between need-persistent (perseverative) and ego-defensive (repressive) reactions to frustration. To this end he attempted to control "the external environment so as to reproduce with two groups of 30 adults each the alternative conditions which had obtained with the children . . ." (17, p. 65). In one group, the so-called informal (non-threatened) group, he intended to arouse need-persistent tensions alone. In the other group, the so-called formal (threatened) group, he intended to arouse need-persistent and ego-defensive reactions. Both groups were tested individually in an interrupted task set-up. The experimental tasks consisted of a set of jigsaw puzzles. The instructions to the informal group clearly put the emphasis on the puzzles rather than on *S*'s performance on the puzzles: *S* was told that *E* was constructing some puzzles for a future experiment and that he needed to know how they would work. Moreover, *S* was informed in advance that work on a puzzle might be interrupted as soon as *E* got the necessary information. Every attempt was made to create working conditions which

⁵ Sanford (19) and Sanford and Risser (20) have questioned the generality of the Rosenzweig and Mason findings, particularly with respect to the mental age relationships they report. Using a group of normal, private school children whose mental age levels "begin near the place where those of Rosenzweig's subjects leave off," Sanford (19, p. 240) reports that the superior recall of incomplete tasks is associated with high intellectual maturity, not with immaturity. To recall one's failures, says Sanford, is the mature response; to forget them, the immature response.

While Sanford touches upon an important problem here, the composition of the sample, the fact that neither his instructions nor his conditions exactly duplicated those of Rosenzweig and Mason must be taken into account. Sanford instructed his Ss to "see how well you can do"; the Rosenzweig and Mason Ss were competing for a prize. Sanford's Ss were permitted to complete one more puzzle than they left incomplete; Rosenzweig and Mason provided for an equal number of each type. It is the former difference which is the more important: Involvement is undoubtedly related to maturity factors, as Sanford points out, but it is also related to the nature of the experimental instructions. The immature *S* may be less ego-involved than the mature *S*, given the same instructions, because he fails to recognize the seriousness of the ego-threat. On the other hand, the immature *S* may feel even more threatened than the mature *S*, given the same instructions, because he is less capable of meeting threat successfully. The experiments in question do not help us choose between these alternatives since the ego-orientation was not identical for the two sets of Ss.

⁶ The average *P* value for these Ss was found to be + 2.75.

would arouse comparatively little tension. The opposite conditions prevailed for the formal group. Here *S* was told that the tasks he would perform constituted an intelligence test. By a variety of procedural devices (see 17, p. 66), the self-esteem threat was further made explicit. For example, a stop watch was now clearly in evidence. Moreover, *S* was given to understand that, if he did not complete a given puzzle within the allotted time, he had failed to score on that task.

It is important to emphasize that Rosenzweig was now attempting to impose from without differences which in the earlier experiment pertained by virtue of the personality structure of the *Ss*. In other words, in Alper's terminology, Rosenzweig expected the formal group to be ego-involved, the informal group to be merely task-involved. Following through the tentative leads from the Rosenzweig and Mason experiment, Rosenzweig hypothesized that the formal group would show a reversed Zeigarnik ratio, i.e., superior recall of completed tasks, the informal group the usual Zeigarnik effect, i.e., superior recall of incomplete tasks.

Since not all *Ss* were given the same number of tasks, percentage scores were again necessary. An intrasession recall index, *R. I.*, was computed for each *S* as follows:

$$100 \frac{(\text{completed} - \text{incomplete tasks recalled})}{(\text{completed} + \text{incomplete tasks recalled})} = \text{R. I.}$$

To support the hypotheses the informal group should show a negative *R. I.*, the formal group, a positive *R. I.* The results do support the hypotheses: Rosenzweig reports a mean index for the informal session of -7.65 ; for the formal session, $+2.95$. The significance of the difference was not tested.

Inspection of the data revealed three recall groups in each session. To take

this fact into consideration, Rosenzweig tabulated for each session the number of *Ss* who recalled a preponderance of incomplete tasks, and the number who recalled an equal number of both types of tasks. Analysis of these data by means of the X^2 technique yielded a X^2 value of 8.64. This value was significant beyond the five per cent level of confidence. Rosenzweig concluded that the preponderant direction of selective recall depends upon the situation in which *S* is placed: The Zeigarnik effect holds for non-stressful conditions, but it does not hold when *S* experiences the situation as one in which he may not show up too well, i.e., one which threatens his self-esteem.

We can now consider Glixman's criticisms of Rosenzweig's study. As noted earlier, Glixman objects to the measure of recall used by Rosenzweig. "The point of view taken in the present paper," says Glixman (9, p. 494), "is that both the ratio described and the comparison of recall of completed with the recall of incomplete tasks upon which are based the values of X^2 are inadequate. So long as recall of incomplete and completed tasks are compared with each other, there is no way of telling whether the change in ratios or the X^2 cited is attributable to change in recall of incomplete tasks, recall of completed tasks, or both."

Now Glixman's statement about "change," taken at face value, might be important for the theory of selective recall if one were permitted to report merely the final statistic (in this case, the mean recall difference score for each session, or the X^2 value) without reference to the original data. But since the investigator must always refer back to his original data in order to interpret his

final statistic, even if his readers do not, the significance of Glixman's objection is not clear. Moreover, even if in Rosenzweig's experiment the reader had available *only* the final intrasession recall difference scores, the scores for the two sessions would indeed reveal the pertinent fact—the preponderant *direction* of recall tends to be a function of experimental setting. The *particular* values in the formula are a function of the particular sample used. The final relationship revealed that the direction of the recall index, whether positive or negative, is the vital datum for purposes of generalization.

What Glixman would have Rosenzweig do is to test hypotheses which interested Glixman, not Rosenzweig; namely, (1) the effect of increased stress on the recall of incompleting tasks, and (2) the effect of increased stress on the recall of completed tasks. Given Rosenzweig's design, Rosenzweig could have tested these hypotheses had he *not* used the interrupted task method as a means of creating "stress" experiences. The interrupted task technique imposes on *S* within a given experimental hour two experiences: Some tasks are completed, others are not completed. Thus, completion is experienced within a partial setting of incompletion, and incompletion in a partial setting of completion. The constellation of forces operating on *S* include *both* completion and incompletion. Under these circumstances, *S*'s recall of completed tasks is psychologically not independent of his recall of incompleting tasks, and vice versa. To insist on separating the two types of recall, then, as Glixman does in using an intersession measure, is to disregard the very experience which *E* has imposed on *S* experimentally. The "confounding" which Glixman claims he wishes to avoid by the intersession measure (9, p. 452) is exactly what he should be measuring when the interrupted task technique is

used. Since Glixman has taken the position that "a mistaken emphasis upon 'the individual as an individual' appears to be one of the reasons for the errors committed by Rosenzweig and Alper" (9, p. 505), he would perhaps argue that the writer should not impose "the individual as an individual" on other *Es*. But this is not the issue here. Whether *E* is investigating group or individual trends does not alter the fact that in order to understand behavior, the psychologist must use scores which can reflect the field forces which he has at least wittingly created for *S*. Regardless of the research orientation, then, when the interrupted task method is used, an intrasession measure of recall is correct; an intersession measure is inadequate.

Glixman makes much of the point that the interpretation of the data differs, depending on the recall score used. What happens to Rosenzweig's data when the intersession score is applied to them? Glixman finds that the difference between means for the recall of incompleting tasks in the two sessions is not statistically significant ($t = 0.23$; $P = 0.82$), while the difference between means for the recall of completed tasks in the two sessions approaches significance ($t = 1.90$; $P = 0.06$). Since the recall of completed tasks tends to be greater under stress than under non-stress, Glixman is willing to go along with Rosenzweig's original interpretation of his data: Under stress the ego is more likely to come to its defense by recalling more successes. Glixman is less ready, however, to accept Rosenzweig's handling of incompleting tasks. "If the recall change for incompleting tasks is considered," says Glixman (9, p. 496), "Rosenzweig seems to have extracted two countervailing tendencies from thin air; on the basis of no change in recall of incompleting tasks, he has suggested that there are coexistent tend-

encies to recall and to forget incompleting activities." The intersession measure for incompleting tasks, indeed, permits no statement about differences in the recall of incompleting tasks as a function of stress, since the null hypothesis which Glixman wishes to test cannot be rejected. But Rosenzweig's statements about the recall of incompleting tasks are not based on intersession measures. He was concerned with the relative recall of completed and incompleting tasks within a given experimental condition. Had he been concerned to test Glixman's intersession hypotheses, a more adequate design would have provided for the testing of one group of Ss who experience only incompleting under non-stress, and of another group of Ss who experience only incompleting under stress. A third group of Ss exposed only to completed tasks under non-stress would have to be compared with a fourth group of Ss who also experience only completed tasks under stress. To test Glixman's hypotheses by using the Zeigarnik interrupted task technique is to *introduce* serious methodological ambiguities rather than to avoid them.

Long before Glixman had voiced his objections to Rosenzweig's study, it was clear that research in the area of selective recall had raised more questions than it had settled. Few investigators could unequivocally reproduce Zeigarnik's (24) findings. Instead, most studies revealed several recall patterns. Rosenzweig and Mason (18), for example, reported four recall patterns; Rosenzweig (17) reported three. No one pattern achieved statistical significance over the others, and no investigator had proposed a theory which covered all of the patterns. Yet one fact was becoming increasingly clear: The preponderant direction of selective recall was not the same for all Ss even when the objective conditions for recall were the

same. This formulation of the problem served as the starting-point for Alper's (1, 2) research. It seemed probable, for example, that previous investigators had failed to discover clear-cut laws of selective recall because they had not been *sufficiently* concerned with the individual recaller. Previous studies had dealt primarily with group recall data only. The few investigators who had suggested positive relationships between selective recall and personality factors, subsequent to the Rosenzweig and Mason (18) study, in most cases based these suggestions either on attempts to relate isolated personality traits to selective recall, or on incidental observations of behavioral differences of Ss who showed different recall patterns. Frequently, personality relationships were mentioned only in footnotes appended to the text, not in the text itself. This literature has been reviewed elsewhere by Alper (1).

Starting with the hypothesis that a given experimental instruction does not necessarily establish the same set in all Ss, that task-orientation does not necessarily assure task-involvement, nor ego-orientation ego-involvement, Alper attacked the problem of selective recall from the point of view of the individual who is doing the recalling. The fact that some Ss recall a preponderance of completed tasks even in objectively non-stressful conditions suggested the possibility that the recall of completed tasks in an informal setting could be an ego-defensive reaction similar to that postulated by Rosenzweig and Mason (18) and Rosenzweig (17) for the recall of completed tasks in a stressful setting. An S with low tolerance for self-esteem threat could presumably feel threatened (ego-involved) even when the instructions are merely task-oriented, whereas an S with high tolerance for self-esteem threat would be merely task-involved under such conditions.

Any given sample of Ss, unselected for personality factors, might be expected to include Ss with low tolerance and Ss with high tolerance. If so, such a sample of Ss should yield no statistically significant differences between the total number of completed and the total number of incomplete tasks recalled if, experimentally, there is an equal number of completed and incomplete tasks to be recalled. This was Alper's (1, p. 405 f.) Hypothesis I, the hypothesis tested in the experiment which Glixman (9) contends was incorrectly analyzed. In the same paper Alper set up two additional hypotheses relating the direction of selective recall to consistent personality factors (Hypothesis II) and predicting that the *same* individual would react to informal and formal conditions in a way which was consistent with his self-esteem needs (Hypothesis III). The data in support of Hypothesis I were reported in the paper which Glixman criticizes. The data in support of Hypotheses II and III were summarized by Alper in that paper and reported in detail in a second paper (2).

Alper's research differed from previous studies in this area in two important respects: (1) The emphasis was placed not only on a study of individuals, but indeed, on the *same* individuals in two different experimental settings; and (2) the selective recall results were to be correlated with extensive personality data.

Ten draft-age male adults served as Ss in Alper's experiment. Each S was tested in an informal session, Session I, and a week later in a formal session, Session II. The instructions for the informal session followed closely Rosenzweig's (17) informal instructions. The instructions for the formal session were patterned on Rosenzweig's "intelligence test" ruse. But now the stress was made even more explicit by having S

work on a set of tasks (scrambled phrases to be assembled into meaningful sentences) in the presence of a more achieving male contemporary and an attractive female "recorder," both of whom were E's "accomplices." Since the Ss, moreover, were told that this "test" was one which the Army was finding useful in selecting candidates for the then wartime Officer Training School, Session II was clearly intended to threaten self-esteem. Ostensibly, this setting would be even more stressful than Rosenzweig's (17) formal condition. The interrupted task technique was used in both sessions.

To test Hypothesis I, Alper used an intrasession recall score. The percentage of completed tasks recalled within a given session was compared with the percentage of incomplete tasks recalled in the same session,⁷ and the differences were analyzed by means of the *t* test, the percentage completed and the percentage incomplete scores having been adjusted for non-linearity. The following *t* values were obtained:

Session I: $t = 1.52$; $P = .20 - .10$,

Session II: $t = 0.65$; $P = .60 - .50$.

These data support Hypothesis I since the differences in selective recall do not reach statistical significance.

Hypotheses II and III were tested by correlating selective recall and personality data. The statistical technique used was syndrome analysis. The personality data were obtained from the intensive forty-hour laboratory study of each S made by twenty-four co-workers at the Harvard Psychological Clinic.⁸

⁷ Percentage, rather than absolute, scores were used in order to adjust for the fact that under stress not all Ss completed the same number of tasks. See Alper (1, p. 412) for further details on scoring.

⁸ For a brief description of the nature of the personality data obtained, see Alper (1, p. 405).

The syndrome analysis supported Hypotheses II and III and provided the basis for setting up a theory of selective recall which permits predictions of the preponderant direction of selective recall given certain personality structures.

Two major syndromes emerged from Alper's study: (1) recall of a preponderance of incompleting tasks in an informal setting and a preponderance of completed tasks in a formal setting; and (2) the recall of a preponderance of completed tasks in an informal setting and of incompleting tasks in a formal setting.⁹ Correlations with personality data identified the former pattern with high tolerance for failure (ego-strength);¹⁰ the latter, with low tolerance for failure (ego-weakness). The direction of selective recall was thus found to be consistent with S's general self-esteem needs.¹¹ The *individual*, in other words, is lawful even though the group may not appear to be so.¹² This

⁹ Two additional patterns of selective recall were suggested by Alper's (2) clinical data: a preponderant recall of completed tasks in both stressful and non-stressful situations, and a preponderant recall of incompleting tasks in both stressful and non-stressful situations. These patterns did not occur with sufficient frequency in the present sample to warrant detailed analysis. Some suggestions based on the clinical data, however, are included in Alper's (2) paper.

¹⁰ The writer wishes to apologize here for an error appearing in the summary of her second article (2, p. 135) which she failed to catch in galley proof. The second paragraph, second sentence, should read: "The recall of completed tasks when self-esteem is objectively threatened is a pattern characteristic of the strong ego who needs to protect his self-esteem only when it is objectively threatened." In the article, the word "completed" in the above sentence unfortunately appears as "incompleted." This, of course, would be inconsistent with the text.

¹¹ For the application of these general principles to a single case, see White, Tomkins and Alper (23).

¹² Postman and Solomon (14, p. 355) suggest that this principle also applies to recognition thresholds.

is essentially the thesis developed recently by Klein and Schlesinger (11) in the area of perception, and by Frenkel-Brunswick (7) for even broader psychological processes.

The two patterns of recall suggested to Alper the following dynamics: Since it is the S with high tolerance for failure who recalls more incompleting tasks in the informal session, recall of such tasks was equated with the Zeigarnik effect, i.e., with task tensions. Such tensions can function readily if S is not under constant need to defend his self-esteem. When the situation is *objectively* threatening, however, this S does defend his self-esteem. To recall completed tasks under these circumstances is to focus on one's successes at the expense of one's failures. The weak ego S, on the other hand, needs to defend self-esteem even when the objective situation is not threatening. This he does in the informal session by recalling completed tasks, which are for him in this context dynamically equivalent to successes. When the objective stress is high, however, his defenses no longer operate efficiently. Overwhelmed by the failure experience, he seems to need to admit his failures before others accuse him of failing. This can be termed an ego-offensive mechanism and is essentially the defense of a weak, insecure individual. Thus the recall of incompleting tasks in an informal setting is not psychologically equivalent to the recall of incompleting tasks in a stressful setting. Only the former corresponds to the Zeigarnik effect.

If these relationships between selective recall and personality structure hold, and they seem to, as shown by Alper (3) in a subsequent prediction experiment, then future studies which use Ss unselected for personality structure will continue to yield diverse or equivocal findings. The personality needs of the Ss, and not merely the ex-

ternal conditions set up by *E*, determine the direction of selective recall. In this connection it is interesting to note that when Rosenzweig's (17, Table 1, p. 68) data are analyzed by means of the *t* test, as were Alper's (1) data, the difference in recall of completed and incomplete tasks within each session is not statistically significant. The *t* value for the informal session is 1.84; the *P* value, .10 - .05. The *t* value for the formal session is 0.52; the *P*, .70 - .60. The predominant recall tendency in the informal session, moreover, is in favor of completed tasks. This was true for Alper's Session I. It is also true for Glaxman's (10) Situation I, his least stressful setting, since approximately 63 per cent of the completed tasks are recalled, as contrasted with 55 per cent of the incomplete tasks.

Indeed, it can be shown that Glaxman's data are in complete agreement with the Rosenzweig and Alper findings. Intrasession comparisons of the mean percentage of incomplete tasks recalled within a given session with the mean percentage of completed tasks recalled in that same session, the computations being based on Glaxman's data (10, Table I, p. 289 and Table IV, p. 291) and analyzed by means of Zubin's (25) nomographs, yield *P* values for each of

Glaxman's three situations which do *not* approach the 5 per cent level of confidence. These data are summarized below in Table I. Thus, the intrasession analysis of Rosenzweig's, Alper's and Glaxman's group data yields *identical*, not diverse selective recall data! Since these analyses deal only with group trends, this finding is consistent with what Alper had hypothesized in her own study (1). Without personality data, however, the analysis of all three experiments would have to stop at this level and the dynamics of selective recall would still elude the investigator.¹³

This was the point that Alper was attempting to establish. Glaxman (9, p. 506) argues, however, that both "Alper

¹³ The writer had discussed with Dr. Rosenzweig the application of this theory to the results of the Rosenzweig and Mason (18) study as long ago as 1943. At that time Dr. Rosenzweig agreed that the theory might also apply to his study. The preponderant recall of incomplete tasks in a stressful session, he felt, might indeed occur because pride was too readily wounded rather than because pride had not been wounded at all. The fact that the *Ss* were crippled children whose need for praise and success might be expected to be even higher than that of normal, well-adjusted children, was a point which strongly favored the application of the theory to the Rosenzweig and Mason (18) results.

TABLE I

MEAN PER CENT OF COMPLETED AND OF INCOMPLETED TASKS RECALLED
UNDER THREE SITUATIONS WHICH INCREASE IN STRESSFULNESS

Data adapted from Glaxman (10, Table I, p. 289, and Table IV, p. 291), and analyzed by means of Zubin's (25) nomographs.

	Form A			Form B		
	I	C	P	I	C	P
Situation I	.56	.65	>0.05	.54	.61	>0.05
Situation II	.49	.67	>0.05	.48	.58	>0.05
Situation III	.49	.51	>0.05	.38	.53	>0.05

I refers to mean per cent of incomplete tasks recalled.

C refers to mean per cent of completed tasks recalled.

and Rosenzweig were confused about the purposes of their experiment." "If the experiments were intended for the study of the relationship between recall changes and personality factors, the experimental designs should have included a personality classification of the subjects. Since there were no personality classifications, a safe assumption is that the purpose of the experiments was to study recall changes as a function of stress. For this purpose, stated in simple form, the factors which make for individual differences are largely irrelevant." Glixman (9, p. 500) restates Alper's Hypothesis I, therefore, as follows: "If recall changes for incompleting and completed tasks are largely a function of personality, and if a sample of subjects is selected randomly with respect to personality variables, then there should be no significant differences when either kind of recall in a stress situation is compared with the corresponding recall in a non-stress situation." To test this hypothesis, Glixman applies his intersession measure to Alper's original recall data, reporting a "near-significant decrease" of incompleting tasks as stress increases ($t = 2.04$; $P = 0.07$), and a "significant decrease" of completed tasks as stress increases ($t = 2.84$; $P = 0.01$).

But this writer would raise the same objections to applying Glixman's intersession measure to Alper's data as she raised to applying it to the Rosenzweig data; namely, (1) the intersession measure is unsuitable since it disregards the psychological interdependence of completion and incompleting which the interrupted task method imposes on *S*; and (2) the hypotheses which Glixman wishes to test require a different experimental design. Moreover, to have included "a personality classification" of the *Ss* in advance of any experimental findings with respect to the role of specific personality factors, as Glixman

would have had Rosenzweig and Alper do, would have been premature. The classification which Glixman requests is, of course, included in Alper's second paper (2), on the basis of the syndrome analysis. The fact that this analysis had been made for these *Ss* was included in Alper's first paper (1), along with a brief summary of the personality findings.

In addition to criticizing Alper's recall measure, Glixman objects to the fact that only ten *Ss* were used in Alper's study. The sample, he feels, was too small. He recognizes, however, that since "her subjects appeared in *both* situations, the efficiency of her experiment is greater than is indicated by the small number of subjects" (10, p. 283). A fundamental point which Glixman overlooks is the manner in which clinical data are used in Alper's study. Glixman (10, p. 283) complains, "Since there is an investment of 400 hours of clinical examination (1, p. 405, footnote 3) which will be used to account for the results based on only 10 subjects—no matter how efficiently used—the implied criticism seems justified." As already noted, in Alper's experiment the personality data were based on an intensive forty-hour clinical study of each *S*. Thus, any increase in sample size would necessarily have increased the number of hours invested, not only to the extent of 40 additional hours for each additional *S* but actually countless additional hours for each of the twenty-four clinicians individually and for the twenty-four clinical co-workers as a group. Clinical studies *are* time-consuming. Yet there is, as yet, no other way to obtain the necessary personality data. Clinical studies, therefore, must begin with small samples and be analyzed by means of small sample statistics. Only afterwards can the leads which such studies disclose be tested on

the larger samples which make certain non-clinicians more comfortable.

In designing his own experiment, Glixman is explicit in wanting to avoid "the errors" of previous researchers, including "a mistaken emphasis upon 'the individual as an individual'" (9, p. 505). "To determine the effects of increasing threat to self-esteem on recall of completed and incompleted activities" (10, p. 289) he used 120 Ss (college students) in an elaborate experimental design based on a three-way classification (Stress situation \times Sequence of task presentation \times Test form). Relying on the complexity of the design and on the large total number of Ss, Glixman confined the number of Ss in each cell to 5. Three situations which differed in objective stressfulness were used. In each situation, the tasks (20 heterogeneous paper and pencil problems) were presented within the interrupted task framework to different groups of Ss in the competitive atmosphere of the college classroom. Selective recall as well as resumption of incompleted tasks was measured.

In Situation I, labeled "neutral," Glixman intended, apparently, merely to task-orient the Ss. He recognizes, however, that "in the competitive atmosphere of the classroom it is almost impossible to make an S feel that his ability—or lack of it—is not responsible for his performance. . . . The 'neutral' situation, then, should be regarded only as less stressful than the other situations" (10, footnote 5, p. 284). Situation II emphasized that S's performance would be taken as a measure of his ability, the implication being that incompletion constituted failure. The atmosphere here was clearly more stressful than Situation I since the instructions stated that the tasks could be used as a "test" for weeding out students who were not likely to have successful college careers. The Situation

II Ss were told that their performance would be used for establishing norms for the rest of the University. Speed and accuracy were emphasized. In Situation III, the ego-orientation was even more explicit. These Ss were told not only what the "test" tested but also that *they* were being tested. Clearly, most students would consider it important to do well.¹⁴

Glixman's intersession recall results can be summarized briefly as follows: As stress increases, significantly fewer incompleted tasks are recalled; as stress increases, the decrease in the recall of completed activities is not significant.

For purposes of theorizing and discussion, Glixman (10, p. 292) brings together the intersession recall results of the Rosenzweig, Alper, and Glixman experiments and summarizes them as reproduced below:

Recall of incompleted activities as stress increases

Rosenzweig: Non-significant decrease ($t = 0.23$; $P = 0.82$).

Alper: Near-significant decrease ($t = 2.04$; $P = 0.07$).

Glixman: Significant decrease ($F = 4.38$; $P < 0.05$).

Recall of completed activities as stress increases

Rosenzweig: Near-significant increase ($t = 1.90$; $P = 0.06$).

Alper: Significant decrease ($t = 2.84$; $P = 0.01$).

Glixman: Non-significant decrease ($F = 0.17$; $P > 0.05$).

On the basis of these data, Glixman (10, p. 292) concludes: (1) his study is the

¹⁴ Glixman (10, p. 288) notes that "in the first two situations, most of the Ss did not indicate name or sex on their data sheets." This is an important point in evaluating the "stressfulness" of the situation. It is regrettable, therefore, that complete uniformity with respect to anonymity was not maintained.

first to yield a significant decrease in the recall of incompleting activities as a function of threat to self-esteem; and (2) "Alper's finding with respect to completed activities is markedly unusual; it is totally unexpected in the context in which it appeared. There seems to be no obvious reason for the recall of *completed* activities to decrease as stress increases."

To reconcile the intersession findings in the three studies, Glixman makes three assumptions. These assumptions will be examined briefly. His first assumption is that Alper's "completed activities" did not produce "feelings of completion." (In Alper's experiment, Ss were told that each scrambled sentence could be assembled in four different ways. The solvable sentences were so constructed, however, that Ss typically achieved no more than two solutions per sentence within the allotted time in both the stressful and non-stressful sessions.)¹⁵ Glixman is inclined to believe that the attainment of some solutions, when more are possible, is more "involving" than the attainment of no solutions at all (10, p. 293). By more "involving" he presumably means that the Ss experienced the "completed activities" as even more incomplete than the "incompleting activities." His argument, however, is not borne out by Alper's experience with these Ss: many

Ss stated that the attainment of even a single solution for a given sentence under the stressful conditions *was* a "success" experience. This was particularly characteristic of the strong ego Ss, the very Ss who recalled more completed than incompleting tasks under stress. The writer would question, therefore, the validity of Glixman's assumption. Further research is indeed necessary on this point, however, as Glixman has also suggested (10, p. 293).

Glixman's second assumption concerns the relative stressfulness of the situations in the three experiments. Rosenzweig's stress situation, he feels, was not as stressful as either Alper's or Glixman's: "Rosenzweig's situation falls below an unspecified critical point, and the others fall beyond the critical point" (10, p. 293). This assumption is, of course, vague and highly speculative. The investigator who wishes to test it would have little to guide him.

Glixman's third assumption is that "increased recall of completed activities is a more superficial defense mechanism than is decreased recall of incompleting activities, and as stress is increased beyond a critical point the increase in recall will disappear. If the situation is threatening enough, a decrease in recall of completed activities may result" (10, p. 293). Logically, then, Glixman would have to claim that Alper's stress situation was even more stressful than Glixman's Situation III. Yet this seems not to have been Glixman's intention in designing his experiment. Nor would such a claim be consistent with his inability in Alper's study to find "an obvious reason for the recall of *completed* activities to decrease as stress increases" (10, p. 292). The deterioration of performance under extreme stress is now a well-known psychological fact. Thus, if Alper's stress situation was more stressful than Glixman's, a decrease in

¹⁵ The device of multiple alternative solutions had been introduced by Alper in order to assure that all Ss could be kept at each task for the same amount of time. The fact that S does not achieve as many alternative solutions as a contemporary who is presumably working under identical conditions would, of course, tend to increase ego-involvement. This is one of the advantages of the multiple alternative solutions device within the context of Alper's design. Glixman may be correct in claiming, however, that the attainment of a single solution under these circumstances does not constitute "success." But if this were so, then Alper's stressful session would be all the more "stressful" for S.

recall of completed activities as stress increases *should* occur. But then the "near-significant" decrease in the recall of incompleted activities in the intersession analysis of Alper's data would be difficult to account for. This decrease would also have to be statistically significant. As the intersession data stand, they do not consistently support Glixman's theory.

The theory as proposed is both highly speculative and not easy to follow. It is not surprising, therefore, that the reviewers are unconvinced. Glixman does not make a good case for the lawfulness of selective recall in spite of the statistically significant findings he reports. For resumption, he can make no case at all. None of his resumption findings is statistically significant. It would appear that an elaborate experimental design cannot serve as a substitute for clinical data in an area where the meaning of the situation to the individual *S* may be the crucial variable.

The final conclusion which the reader would draw from the reviewers' treatments of the three studies, the unsuitability of the interrupted task technique for studying selective recall, remains for further discussion. Sears (22, p. 113), accepting Glixman's argument for an intersession recall score, is of the opinion that the interrupted task method is cumbersome in that it "introduces an additional variable—the Zeigarnik effect—that is already known to influence recall, the very process that serves as the dependent variable in the experiments." The only advantage of the method, he feels, is "that it provides a means of making subjects fail." Since there are other methods for doing this, he favors discarding the interrupted task technique in future studies of the effects of stress on selective recall. He states unequivocally, "When a research operation requires as much discussion of its 'psychological meaning' as inter-

ruption does, it is time to find a new operation." He goes on to say, however, "This criticism applies to the use of interruption as a procedure for securing failure reactions in the study of repression or other memorial functions. It does not necessarily apply to the use made of it in such a study as that of Alper (22) in which the subjects' reactions to completed and incompleted tasks are used as measures in a battery of test situations for the discovery of trait congruences" (22, p. 113 f.). Sears is referring here to Alper's second paper (2). Since he finds Alper's use of the technique acceptable, Sears would seem not to be objecting to the method *qua* method but rather to the experimental designs of most of the other studies in this area. If the objection centered entirely on the matter of experimental design, Alper would have to agree with Sears, since she has raised similar objections elsewhere (1). In stating that the Zeigarnik effect introduces an additional, intrusive variable, however, Sears's objections go beyond questions of design. He is objecting to the interrupted task-method *qua* method.

The answer to Sears's objections is contained in the recapitulation of Alper's theory, summarized above. The arguments in that theory pertinent to the present problem can be restated briefly, as follows: The Zeigarnik effect is not dynamically equivalent to the recall of incompleted tasks under objective stress. Only if it were equivalent, might it be considered an "additional" variable in stress experiments. The recall of more incompleted than completed tasks under objectively non-stressful conditions (the Zeigarnik effect) is a pattern which characterizes the *S* whose self-esteem needs are not readily threatened. This is the Strong Ego pattern of recall under non-stressful conditions. The recall

of more incompleting than completing tasks under objectively stressful conditions, however, is a pattern which is characteristic of a very different personality structure; namely, the *S* whose ego defenses are too weak to permit him to counteract failure constructively. This is the Weak Ego pattern of recall under stress. Under stress this *S* focusses on his failures, admitting them before others accuse him of having failed.

The preponderant recall of incompleting tasks under stress, therefore, represents not the resolution of task tensions, i.e., not the Zeigarnik effect, but rather an ego-defensive mechanism characteristic of the Weak Ego *S* who protects himself against the overwhelming effects of failure by taking the offensive—at least admitting his failures before others accuse him of having failed. Neither the objective conditions for recall, nor the dynamics of this recall pattern fit the original conditions investigated by Zeigarnik (24). The term "Zeigarnik effect" fits only the preponderant recall of incompleting tasks under non-stressful conditions.¹⁶

The theory of selective recall proposed by Alper can bring into essential agreement the experimental results of Zeigarnik (24), Rosenzweig (17), and Alper (1, 2).¹⁷ There is nothing in the

theory which Glixman's (10) data would preclude.

SUMMARY AND CONCLUSIONS

Recently Sears and Cameron, in separate review articles in the *Annual Review of Psychology* (22), challenged the usefulness of the interrupted task method for the study of selective recall. The reviewers were critical largely be-

lieved above could apply to the results of the other investigators are noted below.

(1) The recall pattern of Zeigarnik's "ambitious" *Ss*, who recalled even more incompleting tasks than did other *Ss* in Zeigarnik's informal setting, seems to fit Alper's strong ego syndrome.

(2) The preponderant recall of incompleting tasks by Zeigarnik's *Ss*, as contrasted with the tendency for a preponderant recall of completing tasks in the informal setting of the Rosenzweig, Alper and Glixman studies, may be a function of the experimental tasks and/or the nature of the "informal" settings in the different experiments. Judging from the results of a study by Alper and Black (5), most of Zeigarnik's tasks would not, in and of themselves, be likely to threaten self-esteem, if left incompleting. But, puzzles (Rosenzweig), scrambled phrases (Alper), paper and pencil problems (Glixman), are tasks which could threaten self-esteem, if left incompleting, since they so closely resemble intelligence test tasks. Indeed, it was for this very reason that Alper chose scrambled phrases for the experimental tasks—the tasks had to be acceptable to the *Ss* as an intelligence test in the stressful session. Again, Glixman's warning that his "informal" session was, at best, only less stressful than his other sessions is perhaps pertinent for the Rosenzweig and Alper studies too. In Alper's study, for example, the *Ss* were investing forty hours of their time. They knew that the investigators were engaged in a long-term study of college students. It may well be, therefore, that even the *Ss* with high tolerance for self-esteem threat could not be as relaxed in Alper's informal session as were Zeigarnik's *Ss*. If so, the fact that more completing than incompleting tasks tend to be recalled in Alper's informal session is consistent with Alper's theory. That this tendency does not reach statistical significance, however, is a crucial factor in support of Alper's theory of the role of personality structure in selective recall.

¹⁶ The writer has purposely avoided the term "repression" in discussing the various recall patterns. The Freudian terminology would be both inappropriate and misleading in the present context since: (1) the ego-threat induced in the experimental situation is not, and should not be, intense enough to constitute the basic conditions for Freudian (8) "repression"; and (2) selective recall, as studied by the interrupted-task technique, involves both selective forgetting of the unpleasant and selective forgetting of the pleasant. To stress only the former, as the term "repression" does, is to neglect an important pattern of behavior.

¹⁷ A few examples of how the theory out-

cause the method appeared to have yielded diverse results in three seemingly similar experiments. The experiments in question are Rosenzweig's (17), Alper's (1), and Glixman's (10). Because the writer disagrees with some of the reviewers' criticisms, as well as with Glixman's (9) analysis of the Rosenzweig and Alper studies, the present paper re-examines the hypotheses, the experimental designs, and the data of the three experiments. This re-examination supports the following conclusions:

(1) The reviewers have relied too heavily on Glixman's (9) analysis of the Rosenzweig and Alper studies, accepting as valid his two major assumptions. These assumptions, stated below, are not tenable.

(a) Glixman assumed that the Rosenzweig, Alper and Glixman experiments "were designed in much the same manner to answer the same question" (9, p. 491). In the present paper it has been shown that the hypotheses of the three experiments were *not* the same and that the experimental designs were also not the same. The differences, moreover, are of considerable theoretical importance.

(b) Glixman assumed that the *intra*-session measure of recall used by Rosenzweig and by Alper was inadequate. He would impose, instead, his *intersession* measure on Rosenzweig's and Alper's original data. In the present paper, it has been shown that an intersession score not only disregards the hypotheses which Rosenzweig and Alper designed their experiments to test, but also that an intersession measure is psychologically indefensible when the interrupted task method is used. An intrasession measure *is* correct in the Rosenzweig and Alper studies. It is also applicable to Glixman's data.

(2) When an intrasession measure of recall is used, the Rosenzweig, Alper,

and Glixman studies yield *identical*, not diverse, findings. Stated in the most general form, the finding is this: Analyses of the group *intrasession* recall data obtained by the three experimenters reveal no statistically significant differences in selective recall within a given experimental setting. This finding supports the hypothesis which Alper (1) had designed her experiment to test.

(3) The selective recall of the individual *S* is lawful. The preponderant direction of selective recall is a function of the personality structure of the *S* and not merely of the objective setting which *E* imposes on *S*. Since any randomly selected group of *S*s is likely to include *S*s with different personality structures, group selective recall data alone are unlikely to reveal the lawfulness of selective recall. In order to reveal this lawfulness, selective recall data must be correlated with extensive, independently obtained, personality data derived from a wide variety of clinical techniques. When this procedure was followed by Alper (2), two major recall patterns were revealed. These patterns were summarized in Alper's (1) first paper and presented in detail in a second paper (2). The preponderant recall of incomplete tasks in a non-stressful setting and of completed tasks in a stressful setting was found to be characteristic of the *S* whose ego structure was strong. This recall pattern was therefore labeled the Strong Ego pattern. The preponderant recall of completed tasks in a non-stressful setting and of incomplete tasks in a stressful setting was characteristic of the *S* whose ego structure was weak. This recall pattern was labeled the Weak Ego pattern. The predictive value of these patterns has been supported in a separate experiment (3).

(4) The interrupted task technique is a suitable technique for studying selective recall.

REFERENCES

1. ALPER, THELMA G. Memory for completed and incomplete tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.*, 1946, 41, 403-420.
2. —. Memory for completed and incomplete tasks as a function of personality: correlation between experimental and personality data. *J. Pers.*, 1948, 17, 104-137.
3. —. Predicting the direction of selective recall on the basis of personality structure. (In preparation.)
4. —. Task-orientation vs. ego-orientation in learning and retention. *Amer. J. Psychol.*, 1946, 59, 236-248.
5. —, & BLACK, A. The effect of instruction, task, and population-sample on mental set. *Amer. J. Psychol.*, 1949, 62, 295-299.
6. BARRETT, D. M. Memory in relation to hedonic tone. *Arch. Psychol.*, 1938, 31, No. 223.
7. FRENKEL-BRUNSWIK, ELSE. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, 18, 108-143.
8. FREUD, S. Repression. In *Collected Papers*. London: Hogarth Press, 1925, 4, pp. 84-97.
9. GLIXMAN, A. F. An analysis of the use of the interruption-technique in experimental studies of "repression." *Psychol. Bull.*, 1948, 45, 491-506.
10. —. Recall of completed and incomplete activities under varying degrees of stress. *J. exp. Psychol.*, 1949, 39, 281-295.
11. KLEIN, G. S., & SCHLESINGER, H. Where is the perceiver in perceptual theory? *J. Pers.*, 1949, 18, 32-47.
12. KOFFKA, K. *Principles of Gestalt psychology*. New York: Harcourt, Brace, 1935.
13. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.
14. POSTMAN, L., & SOLOMON, R. L. Perceptual sensitivity to completed and incomplete tasks. *J. Pers.*, 1950, 19, 347-357.
15. PRENTICE, W. C. H. The interruption of tasks. *PSYCHOL. REV.*, 1944, 51, 329-340.
16. RAPAPORT, D. *Emotions and memory*. Baltimore: Williams and Wilkins, 1942.
17. ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistent and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, 32, 64-74.
18. —, & MASON, G. A. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, 24, 247-265.
19. SANFORD, R. N. Age as a factor in the recall of interrupted tasks. *PSYCHOL. REV.*, 1946, 53, 234-240.
20. —, & RISSER, J. What are the conditions of self-defensive forgetting? *J. Pers.*, 1948, 17, 244-260.
21. SEARS, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Council. Bull.*, 1943, No. 51.
22. STONE, C. P. (Ed.). *Annual Review of Psychology, Vol. I*. Stanford: Annual Reviews, 1950.
23. WHITE, R. W., TOMKINS, S. S., & ALPER, THELMA G. The realistic synthesis: a personality study. *J. abnorm. soc. Psychol.*, 1945, 40, 228-248.
24. ZEIGARNIK, BLUMA. Ueber das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, 9, 1-85.
25. ZUBIN, J. Nomographs for determining the significance of the differences between the frequencies of events in two contrasted series or groups. *J. Amer. Stat. Assoc.*, 1939, 34, 539-544.

[MS. received for early publication October 11, 1951]

THE NATURE OF THE RESPONSE IN DISCRIMINATION LEARNING

BY KENNETH W. SPENCE

State University of Iowa

In a recent article in this JOURNAL Weise and Bitterman (11) raise the question as to what the response is in the typical discrimination learning situation. As they point out, two different conceptions have been put forward. One, proposed by Nissen (7) and the writer (9), holds that the subject learns to orient towards and approach the stimulus complex (path, door, alley, window) containing the positive stimulus cue (white, form, etc.) rather than to approach the stimulus complex containing the negative cue.¹ The other, held by Gulliksen and Wolfe (2), conceives of the problem in terms of learning to turn right to one spatial arrangement of the stimulus cues (e.g., black-white) and to turn left to the other spatial arrangement (e.g., white-black). The latter conception assumes that the subject responds to the total stimulus configuration whereas the former assumes that the subject responds to one part of the stimulus complex without regard *necessarily* to the other.

Admitting that the writer's version of the "approach" interpretation was able to handle certain facts which they believed refuted the Nissen version, Weise and Bitterman suggested that a further test could be provided by comparing the rate of learning two different types of discrimination problems. One of these they designated the *simultaneous* problem. Essentially it is the stand-

ard discrimination problem in which on 50 per cent of the trials one of the cues (e.g., black) is on the left and the other (e.g., white) on the right. On the other half of the trials the positions of the cues are reversed. In the second problem, termed the *successive* problem, only one of the cues is present on each trial. If the apparatus is an alley, both paths are either white or both black. The subject must learn to go right when the white cues are present and left when the alleys are black, or vice versa.

Now this latter problem involves a type of discrimination learning which the present writer has never dealt with in any published article. It represents a situation in which no one of the cue elements S_W , S_B , S_L , S_R is consistently reinforced as compared with the others. Unlike the standard situation in which the excitatory strength of the positive cue (S_W) to evoke the approach response becomes steadily greater than that of the negative cue (S_B), so that eventually response is always made to the stimulus complex (alley, window, etc.) that contains it, no one of the stimulus elements or cues in this successive situation acquires greater excitatory strength than the others. How then, in terms of this theory, can one explain the learning of such a problem? The interpretation that the writer has held for some years and has presented at a number of colloquia is essentially the same as that given by Nissen (6). Instead of referring to it as involving a conditional reaction, however, the writer has referred to it as patterned discrimination. According to this conception

¹ Nissen speaks of avoiding the negative cue. The writer's theory of discrimination learning (9) specifically avoided consideration of the condition of punishing a response to the negative cue, and for that reason does not specify avoidance.

the response of approaching certain patterns (combinations) of the stimulus components is consistently reinforced, whereas response to certain others is not. Thus on half the trials the situation is shown below as in A and half as in B.

- A. $\begin{matrix} S_W \\ S_L \end{matrix} \} \rightarrow R_A (+) \quad \begin{matrix} S_W \\ S_R \end{matrix} \} \rightarrow R_A (-)$
 B. $\begin{matrix} S_B \\ S_L \end{matrix} \} \rightarrow R_A (-) \quad \begin{matrix} S_B \\ S_R \end{matrix} \} \rightarrow R_A (+)$

The stimulus compounds, S_W-S_L and S_B-S_R , acquire excitatory tendencies to the response of approaching, the other two (S_B-S_L and S_W-S_R) do not. We can speak here, as experimenters, of a relation, the brightness, i.e., black or white, in connection with the position or spatial cue. The writer would prefer to identify each as discriminable patterns or stimulus compounds *per se*. Discrimination will involve such patterning according to this conception only when no one of the cue members is systematically reinforced more than the others.

There is nothing new, of course, in this conception of stimulus compounds and their differentiation for, as will be recalled, Pavlov (8) demonstrated that one can set up differential conditioned responses to stimulus compounds involving the same elements in different order. Undoubtedly Hull (4) would interpret these different compounds in terms of his afferent interaction hypothesis, and S_W in combination with the cues from the left alley (and the proprioceptive cues from orienting and looking to the left) would now become S_W' , whereas it would be S_W'' when combined with S_R . According to this conception the subject learns to approach S_W' and not S_W'' , the competing stimulus on the particular trial. Likewise on the alternative trials it learns to approach S_B' and not to approach S_B'' .

It is apparent from the above discussion that the discrimination would be more difficult in the case of this *successive* discrimination problem than in the *simultaneous* problem, for the stimulus compounds to be discriminated on each trial are more similar than in the simultaneous problem. Thus S_W-S_L is more different from S_B-S_R than it is from S_W-S_R . At least on this point everyone seems to be in agreement, as Nissen and Weise and Bitterman all reach the same conclusion. On the other hand, according to Weise and Bitterman the Gulliksen-Wolfe interpretation would predict that the reverse would be the case, i.e., that the *simultaneous* problem would be more difficult.

As there was no evidence available directly comparing the relative difficulty of the two types of situations, Weise and Bitterman conducted an experiment in which they employed a multiple (four-unit) discrimination box. Their data clearly showed that the *simultaneous* problem was the more difficult, and they interpreted the finding as opposed to the interpretations of Nissen and the writer which implied that the *successive* problem would be the more difficult.

The finding of Weise and Bitterman was certainly a surprise to the present writer, for in the course of the last 15 years a number of studies had been run in the Iowa laboratories that had tended to suggest that the *successive* problem was relatively the more difficult of the two situations. Thus the Grice (1) and Smith (10) theses had been run under somewhat similar conditions, the former being a *simultaneous* discrimination and the latter a *successive* one. Whereas Grice's subjects learned in some 20 trials, Smith's took 70 or more. A number of our exploratory experiments had also pointed in the same direction. Furthermore, Lawrence (5, 6) has reported two experiments that permitted comparison of the two learn-

ing situations by the same subjects.² After learning first on the *simultaneous* discrimination problem, the subjects subsequently learned the same successive problem. If transfer affects the results in this instance, it certainly favors the successive problem. And yet an examination of the error curves reveals that the curve reached zero at about 25 trials in the case of the *simultaneous* problem whereas it had not yet reached zero by 85 trials in the *successive* problem. In a subsequent study in which the successive problem was learned first and the simultaneous problem was learned second, the *simultaneous* discrimination was again much easier: The curve reached zero by 18 trials whereas the *successive* problem showed errors still being made at 70 trials. In accord with these results a doctor's thesis just completed along the lines of the Lawrence experiments by Heyman (3) in the Iowa Laboratory showed much quicker learning on the *simultaneous* problem.

Nevertheless the Lawrence and Heyman studies involved certain additional irrelevant stimuli that might have been in some way responsible for the disagreement of their findings with those of Weise and Bitterman. Accordingly we instituted an experiment in which an elevated T-maze was employed with black and white cues. It is not necessary to go into detail as to the apparatus and procedure except to state that (1) the maze was a simple T in which the stem was an intermediate gray and the arms of the T were white or black; (2) the non-correction method was used, as this is the procedure specified in the writer's theory (the correction procedure is too complicated); (3) the trials were spaced with a minimum of about five

minutes occurring between trials; and (4) the subjects were hooded rats.³

The results are shown in Table 1, which presents the mean number of errors and trials required in the *simultaneous* and *successive* situations. The finding, it is readily apparent, points in a direction diametrically opposed to that of Weise and Bitterman and is in line with our previous exploratory studies and with the data of Lawrence and Heyman.

TABLE 1
DATA COMPARING LEARNING OF
Simultaneous AND *Successive*
DISCRIMINATION PROBLEMS

Type of problem	Errors		Trials	
	Mean	σ_M	Mean	σ_M
Successive	53.8	7.54	143.0	16.80
Simultaneous	16.6	1.27	59.0	2.77
Difference	37.2		84.0	
t	4.86		4.94	
P	<.001		<.001	

As for the type of discrimination experiment with which the writer's theory is concerned (non-reinforcement and non-correction), he is inclined to believe that the approach interpretation will work. Just why Weise and Bitterman got opposite results is not clear, as it is difficult to interpret the very complex type of discrimination set-up they employed. The simple discrimination situation is sufficiently difficult to deal with theoretically without adding all of the problems that arise as the result of the serial nature of the multiple discrimination set-up along with the fact that it involves a gradient of reinforcement.

The writer would like to take this opportunity to discuss further the type

² Weise and Bitterman cite the Lawrence studies in connection with the description of the successive problem, but fail surprisingly to point out the opposing nature of their results.

³ The writer wishes to acknowledge the assistance of Henry Loess and George Moeller in conducting the experiment.

of patterning that the Gestalters and Gulliksen and Wolffe are talking about when they say that the subject is responding on each trial in the typical discrimination situation to the total configuration. The present writer does not believe this to be the case for, as was pointed out earlier, in such discrimination situations response to the positive cue member (white, triangle, etc.) is being differentially strengthened over response to the others, and hence no conditioning to a pattern is forced.

In discrimination situations not involving differences of degree in the cue stimulus there is little evidence to support the contention that the response is to such a figure-figure pattern. Thus if an animal has been trained to choose a triangle over a circle to get food, substitution of some other form for the negative stimulus does not interfere in any manner with continued response to the positive stimulus so long as no avoidance to the negative stimulus as result of shock or punishment to negative stimulus is involved. That is to say, changing the nature of the total stimulus pattern does not interfere with the response, which would make it seem reasonable to conclude that the response was not necessarily to a pattern or to one of the cues in relation to the other cue.

According to the theoretical view proposed here, response on the basis of such cue-cue relations, or what might be called *transverse* patterning, would take place in non-articulate organisms, but only under conditions that would not permit learning on the basis of a single reinforced component or some simpler type of cue-position pattern (e.g., the type of patterning involved in the successive problem). An experiment that would require a response to the relation of the two cues (e.g., two figures) in the stimulus complex would involve three different stimulus figures being

presented in pairs. Thus the animal might be required to learn the following discrimination simultaneously.

$$\Delta + \text{vs } \bigcirc -; \square + \text{vs } \triangle -;$$

$$\bigcirc + \text{vs } \square -$$

Or in the field of brightness discrimination:

$$20 + \text{vs } 10 -; 5 + \text{vs } 20 -; 10 + \text{vs } 5 -$$

Successful learning of such problems would require that the animal on each trial respond or not respond (by approaching) to a particular figure depending on what the other figure was. Presumably white rats should be able to learn such a pattern discrimination problem; certainly it should be possible of solution by monkeys or chimpanzees. The learning of this type of problem would require a different set of preliminary receptor-exposure acts than simpler discrimination problems involve and one would expect the first solution to be relatively difficult. Subsequent problems of a similar type, however, should show considerable transfer and be learned with not too great difficulty.

REFERENCES

1. GRICE, G. R. The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.*, 1948, 38, 1-15.
2. GULLIKSEN, H., & WOLFFE, D. A. A theory of learning and transfer. *Psychometrika*, 1938, 3, 127-149.
3. HEYMAN, M. Transfer in discrimination learning following three conditions of initial training. Unpublished Ph.D. thesis, State University of Iowa, 1951.
4. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
5. LAWRENCE, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *J. exp. Psychol.*, 1949, 39, 770-784.
6. —. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *J. exp. Psychol.*, 1950, 40, 175-188.

7. NISSEN, H. W. Description of the learned response in discrimination behavior. *PSYCHOL. REV.*, 1950, **57**, 121-131.
8. PAVLOV, I. P. *Conditioned reflexes* (trans. by G. V. Anrep). Oxford: Oxford Univ. Press, 1927.
9. SPENCE, K. W. The nature of discrimination learning in animals. *PSYCHOL. REV.*, 1936, **43**, 427-449.
10. SMITH, M. P. The stimulus trace gradient in visual discrimination learning. *J. comp. physiol. Psychol.*, 1951, **44**, 154-162.
11. WEISE, P., & BITTERMAN, M. E. Response-selection in discriminative learning. *PSYCHOL. REV.*, 1951, **58**, 185-195.

[MS. received for early publication October 1, 1951]



Classics Among

PSYCHOLOGICAL MONOGRAPHS

Thorndike, E. L. *The Mental Life of the Monkey.* 1899, #15. \$*5.50*

Carr, Harvey. *Visual Illusion of Movement During Eye Closure.* 1905, #31. \$*1.25*

Watson, John B. *Kinaesthetic and Organic Sensations: Their Role in the Reactions of the White Rat to the Maze.* 1907, #33, \$*1.00*

Shepherd, William T. *Some Mental Processes of the Rhesus Monkey.* 1910, #52. \$*.75*

Franz, Shepard Ivory and Lafora, Gonzalo R. *On the Functions of the Cerebrum: The Occipital Lobes.* 1911, #56. \$*1.25*

Fernberger, Samuel W. *On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli.* 1912, #61. \$*1.00*

Boring, Edwin G. *Learning in Dementia Praecox.* 1913, #63. \$*1.00*

Langfeld, Herbert S. *On the Psychophysiology of a Prolonged Fast.* 1914, #71. \$*.75*

Franz, Shepard Ivory. I. *Symptomological Differences Associated with Similar Cerebral Lesions in the Insane.* II. *Variations in Distribution of the Motor Centers.* 1915, #81. \$*1.50*

Peckstein, Louis Augustus. *Whole versus Part Methods in Motor Learning.* 1917, #99. \$*.75*

Kjerstad, Conrad L. *The Form of the Learning Curves for Memory.* 1919, #116. \$*1.35*

Tolman, Edward C. *Retroactive Inhibition as Affected by Conditions of Learning.* 1918, #107. \$*.75*

MANY OF THE EARLY MONOGRAPHS ARE OUT-OF-PRINT.

ONLY A LIMITED QUANTITY OF THE ABOVE
NUMBERS ARE AVAILABLE.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Ave. N.W., Washington 5, D. C.

EARLY PUBLICATION

IN

APA JOURNALS

The policy of accepting articles for immediate publication (providing the editor accepts the article and the author is willing to pay the entire cost of increasing the next available issue by enough pages to add his article to the normal content) is now standard practice for all APA journals except *Psychological Abstracts* and the *American Psychologist*.

The actual charge made to the author includes three items:

1. A basic charge of so much per page. This is the minimum amount that it costs to add an additional page to the journal. For 1951 these costs are:

	PER PAGE
JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY	\$17.00
JOURNAL OF APPLIED PSYCHOLOGY	15.00
JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY..	11.00
JOURNAL OF CONSULTING PSYCHOLOGY	14.00
JOURNAL OF EXPERIMENTAL PSYCHOLOGY	14.00
PSYCHOLOGICAL BULLETIN	17.00
PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED	15.00*
PSYCHOLOGICAL REVIEW	14.00

* Since each Psychological Monograph is printed separately, the author of one handled on an early publication basis can be charged exactly the cost of printing. The figure of \$15.00 is an approximate one; the actual figure will be higher for very short monographs and lower for very long ones. The cost will also vary depending upon the amount of special composition and the illustrations used.

These charges are based upon several factors:

- (a) The greater number of words on a particular journal page, the higher the cost per page. Conversely, the fewer words printed on the page, the lower the cost per page.
- (b) The more copies which must be printed, the higher the cost.
- (c) The more expensive the printer, the higher the cost. Compared to the factors listed above, this is not an important difference in the charges made.

2. The full cost of any cuts or other illustrative material, of special composition for tables, and of author's changes in proof.

3. The full cost of any reprints ordered. (Authors of early-publication articles do not receive any free reprints.)